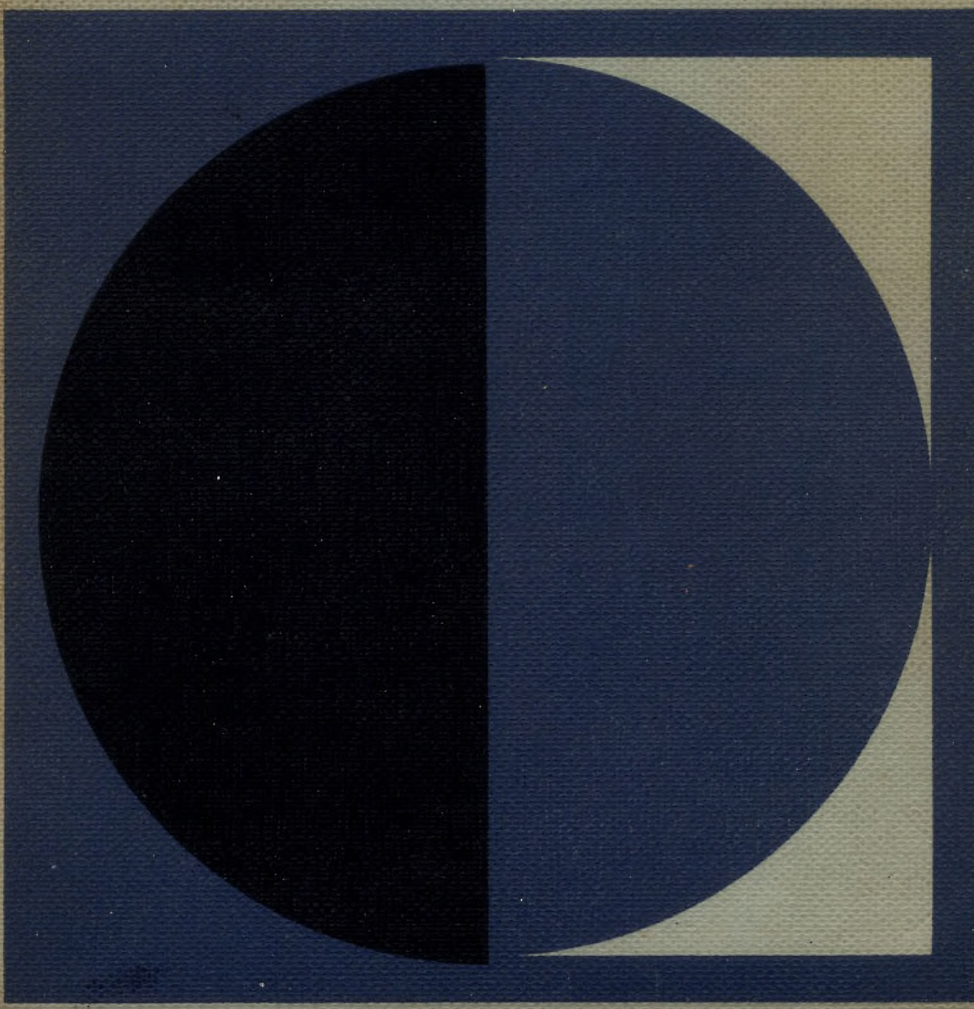
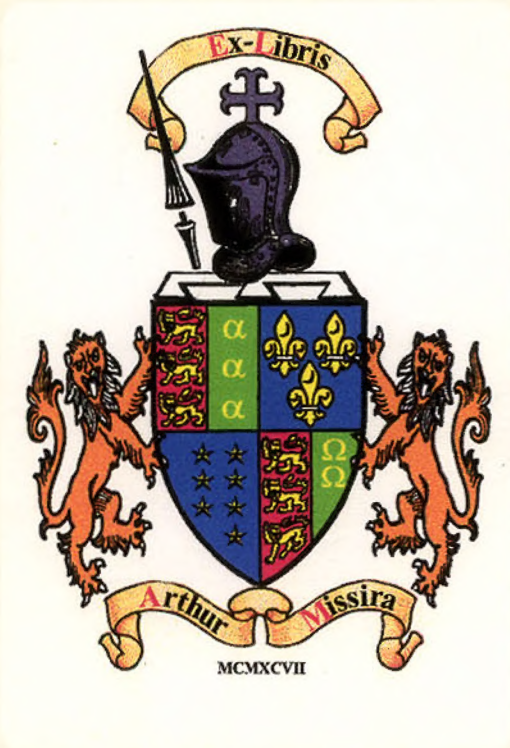
Vectors, Matrices and Linear Equations

H. Neill & A. J. Moakes



Vectors, Matrices & Linear Equations
H. Neill & A.J. Moakes
O&B



Vectors, Matrices and Linear Equations

H. Neill, M.A.
St. Paul's School

A.J.Moakes, M.A., F.I.M.A.
formerly of St. Paul's School

Oliver & Boyd Ltd

OLIVER & BOYD LTD
Tweeddale Court, Edinburgh 1
39A Welbeck Street, London, W.1
First Edition 1967

Printed in Great Britain by
Robert Cunningham and Sons Ltd, Alva

CONTENTS

PREFACE

The trend of pure mathematics in our time is towards the understanding of structure, and at the same time mathematics users are laying great emphasis on linear algebra. The *Mathematics in Education and Industry Project* has accordingly emphasised a structural approach to linear algebra while at the same time showing how the techniques of this subject are applied, especially in geometry of three dimensions and in the solution of simultaneous linear equations.

The course aims to cover the linear algebra and geometry of the M.E.I. pure mathematics syllabus (and, *a fortiori*, of the syllabus in combined subject mathematics) as well as to show the basis of the numerical techniques which are applied to linear equations. It also covers the linear algebra of G.C.E. special papers and of open scholarships; but it does not aim to cover the matrix techniques used in transformation geometry at this level.

We are well aware that in breaking new ground at sixth-form level we may have dug somewhat unevenly, and that there is much to learn about the teaching and testing of this material, but it is hoped that this course will help to meet the needs of those who, like the authors, are engaged in developing it.

We wish to thank all those who in the short time available between trial version and proofs were able to give us the benefit of their comments, notably Mr P. D. Croome.

January 1967 H. N., A. J. M.

ACKNOWLEDGEMENTS

The authors wish to thank the Clarendon Press, Oxford, and Cambridge University Press, Cambridge, for permission to reproduce questions from the mathematical tripos, part I examinations. They would also like to thank various Cambridge colleges for permission to reproduce questions from their scholarship examinations.

PART I

chapter 1

WHAT IS AN ALGEBRA?

1.1 The reader will certainly have met the word algebra, e.g. as a member of the trio 'arithmetic, algebra and geometry'; and the phrases 'Boolean algebra' and 'the algebra of sets' may also be familiar. The full implication of these terms may not, however, be appreciated, viz. that there are *algebras*, in the plural. They must of course have a common character which makes the name applicable to them all. It is the aim of this chapter to bring out what this character is.

One might perhaps expect that what they have in common is the subject-matter: they might all deal with number, but in different ways. This would, however, be far from the case, as we shall show by considering as our first example the algebra of sets. For this purpose we assume that the reader understands the concept of a *set*, how to represent sets by Venn diagrams, and the meaning of *background set*, *subset*, *null set*, *union* and *intersection*. For a full treatment see Moakes: *The Core of Mathematics* (Macmillan).

1.2 In the algebra of sets the background set (or concourse or 'universal' set) is written $\mathscr{E}$. Sets A, B, etc., are subsets of $\mathscr{E}$.†

If we select just A and B for consideration, we can immediately derive from them a further set, also in $\mathscr{E}$, which is written '$A \cap B$', and is called the intersection of A and B. It is represented by the shaded area in the diagram.

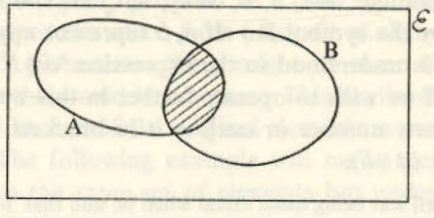
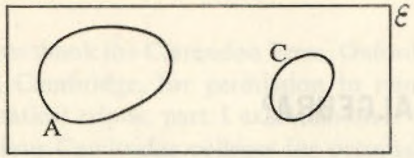


FIG. 1

† Including $\mathscr{E}$ itself and the null set $\varnothing$ as possibilities.

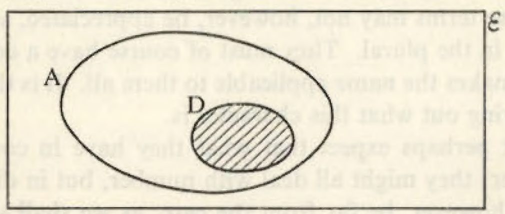
It should be noted that the above diagram illustrates one case only: in particular, if the procedure is always to give us a set of $\mathscr{E}$, then we must have the null set $\varnothing$ defined and understand it to be a subset of $\mathscr{E}$. This figure illustrates a situation in which $A \cap C = \varnothing$:

FIG. 2



With the above proviso, the intersection procedure always gives us a set of $\mathscr{E}$, whatever the original pair we choose; but notice that it does not necessarily give a *new* set, as a further diagram shows:

FIG. 3



For this case $A \cap D = D$; and it will be seen that many other sets, taken in place of A, would give the same result, including $\mathscr{E}$ and D itself, i.e. $D \cap D = D$.

The above is an example of an algebra: it has all the distinguishing marks as we shall see, but it is clear that *(a)* it involves no reference at all to numbers, and *(b)* we do not have to know what the sets are composed of—the sets are simply the units (the algebra's elements) with which we happen to be dealing.†

1.3 Let us now look again for comparison at the familiar algebra in which our elements are the real numbers. (We assume that the reader is sufficiently clear about the nature of this term, embracing all the following and others similar: -2, 3, 0, $-7\frac{1}{2}$, π, $\sqrt{39}$; the whole collection being denoted by the symbol R.) If a, b represent any two of R, then the symbol '+' is understood in the expression '$a+b$' as giving us an element of R. If we wish to operate further in this way we either give a letter to this new number or enclose it in brackets to emphasise its single character: $(a+b)$.

† Bertrand Russell was being quite literal when he said that 'in pure mathematics we do not know what we are talking about'.

So we can, for example, form a new number if desired by combination with x, writing the new number $x+(a+b)$.

We have stated the above procedure with some care because, as the reader is aware, the combination '+' is far from typical. When we use the sign '−', a failure to carry out the correct procedure will give a false result: in fact at the very start even the *order* of the letters a, b in the combination is relevant, giving different results for $(a-b)$, $(b-a)$. The algebra derived from the '−' sign is *non-commutative.*

In case the reader is tempted to suppose that the algebra of sets is less rich that that of numbers (with its four signs $+$, $-$, $\times$, $\div$), we give now two further examples of set-combinations. These and others will be developed in the student's exercises.

(i) $(A-B)$ is defined as the set formed of all those members of A which do not belong to B, for example:

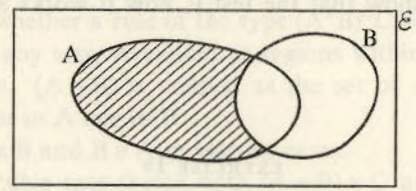


FIG. 4

The reader should verify that in general $A-B \neq B-A$ and that it is possible for $A-B$ to be $\varnothing$ without $A = B$.

(ii) $A \cup B$ is the familiar union of two sets. It is clearly commutative, i.e. AUB and BUA describe the same set.

1.4 A further example of an algebra is that of statements. This takes simple sentences as its elements and combines them by using one or other of certain link-words, e.g. *and, or.* An example will be found in exercise 1a (question 7).

Summarising, we see that an algebra deals with a stated collection of entities (sets, numbers, sentences in our three examples), which are called its *elements.* There must also be a clear and unambiguous *rule of combination,* a process which from any pair of elements will give us another element. (Sometimes there might be more than one law of combination, but this would be a bonus.) Thirdly, the element obtained in this way must be one of the original elements, a provision which is called *closure.* The following example will make these requirements clear. It refers to the same set of elements but under three different laws of combination.

Consider the set of the (unlimited) lines in a plane. Any two lines a and b of this set are taken:

(*i*) a^0b is defined as the point of intersection of the lines.

First, there may be no point—lines may be parallel—but even when it exists *it is not a member of the original set, which consisted of lines.* Closure is not secured, and we have not therefore an algebra.

(*ii*) $a \, ♀ \, b$ is a line bisecting the angle between a, b. If a and b meet we get *two* lines and there is nothing to tell us which one to take. (We also need a definition to cover the cases when they are parallel or coincident.) It is not an algebra, *yet*.

(*iii*) $a \top b$ is a line defined thus, with the help of an axis Ox:

(*a*) If $a \parallel b$, $a \top b$ is $\parallel$ to both and equidistant.

(*b*) Otherwise $a \top b$ is the angle-bisector which has positive or zero gradient with respect to Ox. This is an algebra, not an important one, but it serves to show that the test is how it works and not what its subject-matter is.

EXERCISE 1a

(1) For members of the set R of real numbers we define $a * b = \dfrac{a+b}{2}$. This may be described as the 'split-the-difference' operation. Examine whether this situation satisfies the criteria for an algebra, viz.

(*i*) to have a known set of elements;

(*ii*) to have a combination rule defined for any pair of these elements (including an element combining with itself);

(*iii*) that the result of the combination is also a member of the set.

(2) Test the truth or falsity of the following statements in the algebra of question (1). It is important to note that one case only, chosen to be as simple as possible, can show a statement to be false: but only a general argument can demonstrate the truth of a proposition:

(*i*) $a * b = b * a$ (*ii*) $(a * b) * c = a * (b * c)$ (*iii*) $a * a = a$

(3) Which of the following are algebras? Give reasons.

(*i*) The set of prime numbers under rule '×' (multiplication).

(*ii*) The set of positive integers; under combination θ, where $m \, \theta \, n$ is the L.C.M. of m and n.

(*iii*) The set of numbers of the form $m+n\sqrt{3}$, where m and n are positive integers; under multiplication.

(*iv*) The set $\{0, 1, 2\}$ under:

(*a*) the rule of multiplication 'modulo 3', i.e. from all products $\geqslant 3$, 3 is to be subtracted as many times as possible and the remainder taken as the answer. For example, $2 \times 2 = 1$ (modulo 3).

(*b*) the rule $a * b = \sqrt{a \times b \text{ modulo } 3}$.

(4) Consider the set of all points in a plane. Which of the following are algebras?

(*i*) A $*$ B is the point where the straight line terminated by A and B cuts a fixed line l in the plane.

(*ii*) As (*i*), but the line through AB is unlimited in length.

(*iii*) A^0B is the midpoint of AB.

(*iv*) A.B is the point N such that B is midpoint of AN.

Test in each case whether the combination is commutative, and in cases (*iii*), (*iv*), whether a rule of the type $(A^0B)^0C = A^0(B^0C)$ is true.

(5) A, B, C are any sets, visualised as regions within a rectangle $\mathscr{E}$ in a Venn diagram. (A n B) is defined as the set of all members of $\mathscr{E}$ which are neither in A nor in B.

(*i*) Sketch A n B and B n A in various cases.

(*ii*) For a suitable case sketch both (A n B) n C and A n (B n C).

(*iii*) Simplify A n $\varnothing$, A n $\mathscr{E}$, A n A. (There must in each case be an unambiguous answer, true for all figures, if this is in fact an algebra.)

(6) Simplify:

(*i*) $A - \varnothing$ (*ii*) $A - A$ (*iii*) $A \cup \varnothing$, $A \cup A$, $A \cup \mathscr{E}$

(*iv*) $A \cap \mathscr{E}$, $A \cap A$, $A \cap \varnothing$ (*v*) $(A \cap B) - B$ (*vi*) $(A \cup B) - B$

(7) The connective OR used between statements is understood to mean 'this' or 'that' or both: it is written shortly as $\vee$. The statement p is 'St. Paul's Cathedral is in London' and q is 'President Johnson is in London'. (The former is true and the latter may or may not be true.) Is the statement $p \vee q$ (*a*) necessarily true; (*b*) necessarily false; (*c*) possibly true and possibly false?

Examine also the statements:

p *and* q (also written $p \wedge q$)

not-p $\wedge$ not-q

$(p \wedge q) \vee r$ where r is the statement 'black is white'.

1.5 We have stated the basic properties which constitute an algebra in the broad sense in which we are using the term in this book. If it is an interesting and valuable algebra it will certainly have extra properties,

and the varieties of algebra arising in this way have special names: one type, for example, is called a *group*.

One feature which we look for is whether there is an element in our set which has a special property not shared by others; e.g. in the algebra of sets under intersection, it is uniquely true of $\mathscr{E}$ (which, it will be recalled, is an element of the algebra) that in combination with *any* set X of the system we have $\mathscr{E} \cap X = X$.

Compare, for numbers under addition, the property uniquely held by zero, that $0+x = x$. (Notice that since both systems are commutative we do not have to inquire whether the unique element in question has this property also when placed afterwards.)

An element which has the property just described is called the *identity* element or *neutral* element of the system. It must be emphasised that the element must act in this way when combined with *any* element of the set (and whether placed in front or behind).

EXERCISE 1b

(1) Show that $\varnothing$ is an identity element for sets combined by *union* '$\cup$', but NOT for sets combined by '$-$'.
(2) What is the identity element for numbers combined under multiplication? Is the same element an identity for division?
(3) Show that there is no identity element for sets under *n*, where $A\, n\, B$ = {members of $\mathscr{E}$ which are in neither A nor B}.
(4) The combination $A \Delta B$ for sets A and B is defined as the union of $A-B$ and $B-A$, i.e. it consists of all members of $\mathscr{E}$ which are in A but not in B together with those which are in B but not in A. (It is read as 'delta' and called *symmetric difference.*) Sketch $A \Delta B$ for various cases and show that $\varnothing$ is an identity element.†
Test whether $A \Delta (B \Delta C) = (A \Delta B) \Delta C$.

1.6 A further important property which an algebra may have is that each member may be linked in some obvious way with an 'opposite number'. For example, in the algebra of statements each statement has a *negation* 'not-*p*' (and the negation of not-*p* is *p* again). In the algebra of sets the set A has as an opposite the set $(\mathscr{E}-A)$ of all the other members of $\mathscr{E}$: it is called the *complement* of A and written A'.

† The definition requires it to be unique but we are not asking the reader to prove this.

In important cases the combination of an element and its 'opposite' gives a unique element which we have already met, the neutral or identity element for this combination; if so our vaguely-defined 'opposite' acquires an official status and is called an *inverse*. An important example is among numbers with combination '+': the inverse of x is $(-x)$, the number with reverse sign, because $x+(-x) = 0$ the identity element for the '+' combination.

It is important to note that there is no suggestion of another law (subtraction) in the use of $(-x)$: all the real numbers are signed numbers and our symbol indicates *sign reversal*; e.g. if x is 2 it becomes -2, called negative 2, and vice versa.

EXERCISE 1c

(1) Consider the numbers {0, 2, 4, 6, 8} with combination $\oplus$, addition modulo 10; the units figure only is retained after addition, e.g. $6 \oplus 8 = 4$. Is there an identity element in this system?
Solve the equations (*i*) $2 \oplus x = 0$, (*ii*) $y \oplus 4 = 0$.
Suggest an 'opposite' for 4; is it the inverse for this rule of combination? That is, does $4 \oplus$ (opposite of 4) equal the identity?
(2) Consider the same numbers as in question 1 under $\otimes$ defined as multiplication modulo 10, e.g. $8 \otimes 4 = 2$. By making a full multiplication table show that there is an identity element for this system, but not an inverse for every element.
(3) Show that any *non-zero* number a has an inverse under multiplication, i.e. that $a \times a' = i = a' \times a$ where i is the appropriate identity element. What is i?
(4) Δ signifies 'symmetric difference' of sets, e.g. $A \Delta B$ is shown shaded in the figure. It is defined in exercise 1b, question 4.

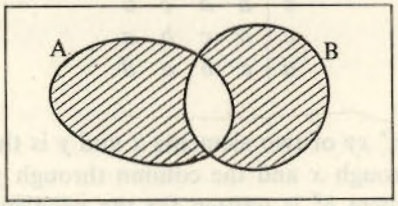


FIG. 5

(*a*) Find a member I of the system, i.e. a set I in $\mathscr{E}$, such that for every set X it is true that $X \Delta I = X$ and $I \Delta X = X$. Give a name to this element.

(*b*) For any member Y of the system, try to obtain a member Z such that $\mathrm{Y} \Delta \mathrm{Z} = \mathrm{I}$. If such a Z exists, what is its relation to Y called?

(5) Show by sketches, which should cover different types of pattern, that

(*i*) $\mathrm{A} \cap (\mathrm{B} \cap \mathrm{C}) = (\mathrm{A} \cap \mathrm{B}) \cap \mathrm{C}$ (*ii*) $\mathrm{A} \cup (\mathrm{B} \cup \mathrm{C}) = (\mathrm{A} \cup \mathrm{B}) \cup \mathrm{C}$

When a relation of this sort is found to hold in general, the system is said to be *associative*.

(6) Consider the set $\{1, 3, 7, 9\}$ under multiplication modulo 10. Have all the elements inverses? Is the system associative?

(7) Show that $\mathscr{E}$ is the neutral element for sets under intersection. Show that if A′ is defined in the usual way (as the complement of A) then A is the complement of A′, but that they are not inverses under intersection: i.e. $\mathrm{A} \cap \mathrm{A}' \neq \mathscr{E}$, the neutral or identity element for this law of combination.

(8) In each of the following examples there is a set of elements and a well-defined law of combination. Consider in each case whether there is (*i*) closure, (*ii*) a neutral element, (*iii*) a unique inverse for each element, (*iv*) commutativity, (*v*) associativity.

(*a*) The set of integers, positive and negative and zero, with the law $a * b = a+b+2$.

(*b*) The set of positive integers with zero, under the law $a \sim b = |a-b|$, the numerical difference.

(*c*) The set of positive integers, under the law $a \circ b = \sqrt{a^2+b^2}$.

(*d*) The set of real numbers, under the law $a \circ b = \sqrt{a^2+b^2}$.

(9) Four elements a, b, c, d are subject to a 'multiplication table'

	a	b	c	d
a	b	a	d	c
b	a	d	c	b
c	d	c	b	a
d	c	b	a	d

where the 'product' xy of two elements x and y is that element which lies in the row through x and the column through y. (For example, $cd = a$, $b^2 = d$, where b^2 is written for the product bb.) Prove that each of the four elements satisfies the relation $x^3 = x$, where x^3 means $x(x^2)$.

Verify that the associative law $x(yz) = (xy)z$ does not always hold, giving two instances of the failure.

A 'product' of the elements a, b, c, d (in that order) is formed by successively multiplying two elements at a time: for example, in the sequences indicated by the groupings $[\{(ab)c\}d]$ or $[(ab)(cd)]$, the order of the letters being unchanged. Find the values of all such 'products'. (*Camb. Schol.* 1964.)

(10) Objects $\ldots \langle -2\rangle, \langle -1\rangle, \langle 0\rangle, \langle 1\rangle, \langle 2\rangle, \ldots$ are given. Two objects $\langle p\rangle$ and $\langle q\rangle$ are *equal* if $p-q$ is a multiple of 4. Prove that:

(*i*) $\langle p\rangle = \langle q\rangle$ and $\langle q\rangle = \langle r\rangle$ imply that $\langle q\rangle = \langle p\rangle$ and $\langle p\rangle = \langle r\rangle$;

(*ii*) each $\langle p\rangle$ is equal to a member of the set S consisting of $\langle 0\rangle$, $\langle 1\rangle$, $\langle 2\rangle$ and $\langle 3\rangle$.

The product $\langle p\rangle * \langle q\rangle$ is defined as $\langle p+q+1\rangle$. Prove that:

(*iii*) $(\langle p\rangle * \langle q\rangle) * \langle r\rangle = \langle p\rangle * (\langle q\rangle * \langle r\rangle)$;

(*iv*) there is a member E of S such that, for every p, $\langle p\rangle * \mathrm{E} = \mathrm{E} * \langle p\rangle = \langle p\rangle$;

(*v*) every member $\langle p\rangle$ of S has an inverse $\langle p\rangle^{-1}$, also a member of S, such that $\langle p\rangle * \langle p\rangle^{-1} = \langle p\rangle^{-1} * \langle p\rangle = \mathrm{E}$.

Complete the multiplication table below:

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$	$\langle 3\rangle$
$\langle 0\rangle$				
$\langle 1\rangle$				
$\langle 2\rangle$				
$\langle 3\rangle$				

Evaluate:

$(\langle 1\rangle * \langle 2\rangle^{-1}) * (\langle 2\rangle * \langle 3\rangle)$ and $\langle 1\rangle * \{\langle 2^{-1}\rangle * (\langle 2\rangle * \langle 1\rangle^{-1})\}$.

(*Camb. Schol.* 1964)

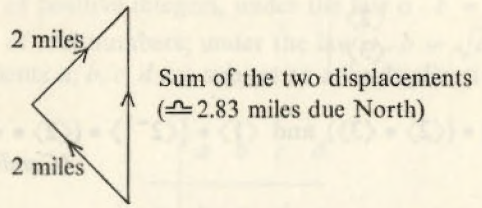
chapter 2

THE PURE ALGEBRA OF VECTORS

2.1 The word VECTOR comes from physics. The physicist distinguishes on one hand quantities which only need a number to specify them fully, and on the other hand quantities which need direction as well. The first are called *scalar* quantities and include temperature; the second include quantities like magnetic field strength, and could be called *directed quantities*.

A simple example of a directed quantity is movement (displacement), e.g. 2 miles in a north-westerly direction. (This is clearly not the same as a movement of 2 miles in a north-easterly direction.) A simple piece of practical geometry shows that we can add displacements. The two displacements we have mentioned, carried out in succession, give us a northerly displacement, of an amount that can be either measured or calculated.

FIG. 6



Directed quantities which can be added in this simple way are called *vector quantities*. Another example of such a quantity is velocity, and there are many more in physics and mathematics.

2.2 When you learned to use numbers of various sorts, as part of pure mathematics, you found that they could be applied in various ways to solve problems in science or in everyday life—but of course only in connection with *scalar quantities* (e.g. a volume of $10\frac{1}{2}$ gallons, a temperature of $-10{\cdot}1°$C).

We shall now develop the mathematics of vectors—as distinct from vector quantities. Like numbers they are completely abstract things; but inevitably we wish to give them a visual embodiment. Just as we

picture the real numbers along a line, so it is useful to see vectors as displacements, at first in a plane and later in three dimensions. Such pictures will help us and will generate ideas, but for definition and for rigorous proof we shall invariably use algebraic methods. For this purpose we shall assume all the properties of *one* algebra, that of real numbers (and when required, subsets such as the integers or the rationals).

EXERCISE 2a

(1) An expedition leaves its base, and its movements on successive days are reckoned to be (*i*) 21 miles on a bearing 080°, (*ii*) 18 miles on a bearing 115°. Find by drawing or calculation (*a*) their final bearing and distance from the base, and (*b*) how far east and how far north or south of it they are.

(2) A man who can row at 4 m.p.h. in still water is rowing steadily, heading due north in a tide which is running westwards at 3 m.p.h. In which direction is he actually travelling and how far will he go in $\frac{1}{2}$ hour?

(3) If the man in question 2 wishes to travel due north, in which direction should he head his boat, and how fast will he travel?

(4) A journey in desert country consists of the following stages: 40 miles due east, 52 miles north-east, and 12 miles on a bearing 337° (i.e. N 23° W). Find (*a*) by drawing to scale, the bearing and distance of the end-point from the starting-point; and (*b*) by calculation, the easterly and northerly components of the total displacement. Hence check the accuracy of your drawing.

(5) An aircraft heads in a direction 300° at an airspeed of 500 knots. Find its true velocity in magnitude and direction if the wind is 50 knots from the S.W.

(6) With the same airspeed and wind velocity as in question 5, the pilot wishes to fly due south. In which direction must he head his aircraft? (Give your result to the nearest degree.) What will be his speed over the ground?

2.3 *Algebraic definition of a vector*

We shall define a vector of (for example) order 3 as a set of three real numbers arranged in order; e.g. in a row, as $(5\ -4,\ 2\frac{1}{2})$, or in a column, as $\begin{pmatrix} 3 \\ 7 \\ -1 \end{pmatrix}$; subject to a law of addition in which corresponding numbers are added together, viz. $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix}$.

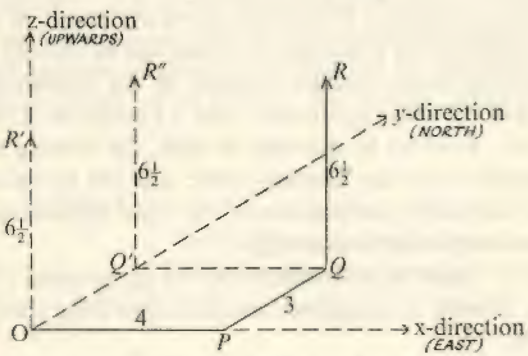
There is no limit to the order of a vector, and in many instances in computing the machine is dealing with vectors of order 100 or more: the law of addition is of the same character, whatever the order.

We shall not at present use two ways of writing vectors; we shall restrict ourselves to the column method except in a few instances where this would be too cumbersome.

The separate numbers which make up a vector are called its *components*.

3-vectors (i.e. vectors of order 3) may be pictured as displacements in space: the first number could represent movement eastward, the second northward and the third upward. The whole vector represents the whole movement, but by virtue of the addition law we can regard it as built up as shown in the diagram, in which the displacement OR (i.e. from O to R) represents the vector $\begin{pmatrix}4\\3\\6\frac{1}{2}\end{pmatrix}$.

FIG. 7



In the diagram the movement $\overline{OP}$ (O to P) or any movement equal and parallel to it, is a picture of the vector $\begin{pmatrix}4\\0\\0\end{pmatrix}$ for which we shall use the single symbol **a**. Similarly $\overline{PQ}$ (or $\overline{OQ'}$ or $\overline{R'R''}$) is $\begin{pmatrix}0\\3\\0\end{pmatrix} = \mathbf{b}$ (say), and $\overline{QR}$ (or $\overline{OR'}$ and $\overline{Q'R''}$) is $\begin{pmatrix}0\\0\\6\frac{1}{2}\end{pmatrix} = \mathbf{c}$. Also $\overline{OR} = \begin{pmatrix}4\\3\\6\frac{1}{2}\end{pmatrix}$, the vector sum of the three.†

† In writing we put a curly line under vector symbols, as $\underset{\sim}{a}$, $\underset{\sim}{0}$.

Notice that for vectors, addition is *commutative*, e.g. $\mathbf{a}+\mathbf{b}$ and $\mathbf{b}+\mathbf{a}$ are both equal to $\begin{pmatrix}4\\3\\0\end{pmatrix}$. See what this means on the diagram.

Also, the addition is associative, i.e. $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$.

The reader should identify each of the *four* vectors in this equation with an appropriate displacement.

It should be noted that here and elsewhere we always draw our set of 3 axes in the same way, described as *right-handed*. The test for this is the corkscrew rule. The handle, turned through 90° from the x-axis to the y-axis, drives the point of the corkscrew in the direction Oz.

2.4 We appear now to have set up an algebra in which the elements are 3-vectors and the rule of composition is known as addition. It appears obvious that the result of 'adding' two 3-vectors is always a 3-vector: but is this always true? Consider the case in 2-vectors:

$$\begin{pmatrix}4\\-2\end{pmatrix}+\begin{pmatrix}-4\\2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

Our rule of closure breaks down unless we are prepared to count the result as a vector (although in geometry it has *no direction*)—an instance where a geometrical definition would have let us down: the algebra is more precise. A suitable single symbol for this zero vector is **0**, to distinguish it clearly from the *number* zero. The context will make clear whether the symbol means the zero 2-vector, as here, or the zero 3-vector.

To simplify diagrams we shall continue to discuss 2-vectors, to show addition properties of such an algebra.

First, $\quad \begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}$ or $\mathbf{a}+\mathbf{0}=\mathbf{a}$.

This shows that the zero is an identity element (and it is clearly unique).

Secondly, our example above, viz. $\begin{pmatrix}4\\-2\end{pmatrix}+\begin{pmatrix}-4\\2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$, clearly shows that we have inverses in this algebra: the two vectors on the left are inverses of each other under vector addition. If we write $\begin{pmatrix}4\\-2\end{pmatrix}=\mathbf{u}$

it seems natural to write $\begin{pmatrix} -4 \\ 2 \end{pmatrix}$ as $(-\mathbf{u})$, read as 'negative **u**'. In general, and for 3-vectors, if **a** is $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ then $-\mathbf{a}$ is $\begin{pmatrix} -a_1 \\ -a_2 \\ -a_3 \end{pmatrix}$.

EXERCISE 2b

(1) **c** and **d** are 2-vectors. Show on a diagram how to form $\mathbf{c}+\mathbf{d}$ and $\mathbf{c}+(-\mathbf{d})$ geometrically; and show that the sum of these is equal to the sum of **c** and itself. (Remember that **c** can be shown in a variety of positions.) What geometrical theorem does the result suggest?

(2) Write down in column form the zero-vector **0** in the system of 3-vectors. Calculate the sum **s** of **a** and **b** where

$$\mathbf{a} = \begin{pmatrix} 5\cdot5 \\ 2\cdot1 \\ 0 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} -8\cdot2 \\ 0 \\ 3\cdot7 \end{pmatrix};$$

and give also the vector $\mathbf{s}'$ which is the inverse of **s**.

Verify that **a** and $(\mathbf{b}+\mathbf{s}')$ are also inverses.

(3) Show with the help of a sketch that addition of displacement vectors is associative, i.e. that $(\mathbf{p}+\mathbf{q})+\mathbf{r} = \mathbf{p}+(\mathbf{q}+\mathbf{r})$.

(*Note*: In the bookwork we have only shown this for a set of vectors parallel to the axes.)

(4) Give two reasons why a sum of money expressed in pounds, shillings and pence cannot be regarded as a vector, although addition is clearly defined.

Matching: an exercise in logic

In many situations one is concerned with matching, e.g. to get correct colour and texture of woollen thread for mending, or to get an efficient substitute for a blown radio valve. We are concerned that the item should be the same as the original in certain essential respects though not necessarily in all. This notion is important in algebra, being called *equivalence*, the criterion for things to be equivalent being defined afresh in every instance. We shall continue to use the word 'match' instead of 'be equivalent to'. These are the essentials for the word to make sense:

(*i*) An item must match itself.

(*ii*) If item x matches item y, then y matches x.

(*iii*) If x matches y and y matches z, then x matches z.

The reader should try this out in some mathematical contexts. Here is a worked example:

Suppose that two positive integers 'match' if their difference is zero or a multiple of 5. Using the symbol $\equiv$, we see at once that

(*i*) $m \equiv m$.

(*ii*) $m \equiv n$ implies $n \equiv m$ e.g. $2 \equiv 7 \Rightarrow 7 \equiv 2$.

(*iii*) $\left.\begin{matrix} m \equiv n \\ n \equiv p \end{matrix}\right\}$ implies $m \equiv p$.

By virtue of this, we only need $\{0, 1, 2, 3, 4\}$ when working modulo 5 because we have here a 'match' for every integer there is.

Matching problems

(*a*) Show that being brother to someone *is not* a matching: being of the same blood-group *is*.

(*b*) In a set of terminated, directed lines in a plane, show that 'being perpendicular to' is not a matching, but 'being parallel to' is, provided you deal in a special way with item (*i*). Is 'being of the same length', leaving direction out of account, a matching? Is 'being parallel *and* the same length' a matching? (If it is so, then as in our example a single symbol such as **a** could be used for a whole mass of equal and parallel directed lines: technically it is called a *representative* of a whole *equivalence* class of displacements in the plane which are equal and parallel to each other.)

(*c*) In a set of triangles, is 'having a side in common with' an equivalence? Is 'being similar to' an equivalence?

2.5 *Vectors in algebra and directed lengths in geometry*

In the light of the idea of matching (equivalence) we see that all the directed lines in space which are equal and parallel to each other (in the same sense) form an equivalence class. When we use an algebraic symbol such as **a**, it stands for the whole class at once—though we may fasten our attention for the moment on only one, in a geometrical diagram.

When we think what the equation $\mathbf{a}+\mathbf{b} = \mathbf{c}$ means in geometry, we have the whole of space filled with directed lines being added together. We readily see that the lines which form the sums will also form an equivalence class, so that there is justification for writing the symbol **c** to represent them all. This logical tidying-up process is discussed again in a note after chapter 8.

chapter 3

VECTORS MULTIPLIED BY NUMBERS

3.1 It seems natural when faced with $\mathbf{a}+\mathbf{a}$ to write it on a simple counting principle as $2\mathbf{a}$, and similarly with any number of vectors. In particular,

$$\begin{pmatrix} 2{\cdot}1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2{\cdot}1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4{\cdot}2 \\ -2 \end{pmatrix} \text{ could also be written as } \begin{pmatrix} 4{\cdot}2 \\ -2 \end{pmatrix} = 2\begin{pmatrix} 2{\cdot}1 \\ -1 \end{pmatrix}$$

By extension we should also write $\begin{pmatrix} 2{\cdot}1 \\ \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 4{\cdot}2 \\ -1 \end{pmatrix}$, or in general, taking a 3-vector for our next example:

$$k\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} ka_1 \\ ka_2 \\ ka_3 \end{pmatrix} \text{ where } k, \text{ like } a_1, a_2 \text{ and } a_3, \text{ is any real number.}$$

(See section 6.4 for a further rigorous discussion of this step.)

Two important special cases occur: first, any vector multiplied by (the number) zero gives the zero *vector*; and secondly,

$$(-1)\mathbf{a} = (-1)\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -a_1 \\ -a_2 \\ -a_3 \end{pmatrix}$$

which we have already met and written as $-\mathbf{a}$, the inverse of $\mathbf{a}$. That this result seems trivial is due to its likeness in appearance to one in the algebra of numbers. It is in fact telling us the *meaning* of multiplying a vector by (-1), viz. a reversal of its direction, to give its inverse.

From this point onwards we can carry out without fuss a process with vectors which we may if we wish call subtraction, but we have no need to regard it as a new combination-method: we form the inverse of the vector and add it. We may however without confusion write a result such as $\mathbf{b}+(-\mathbf{a})$ in the form $\mathbf{b}-\mathbf{a}$ by analogy with arithmetic. Figure 8 gives a useful example.

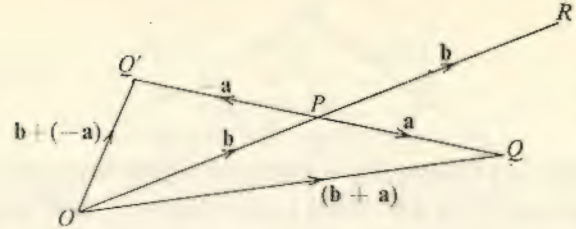


FIG. 8

We can see a number of results on this diagram. Displacement $\overline{Q'P} = \overline{PQ} = \mathbf{a}$, so that:

$$[\mathbf{b}+(-\mathbf{a})]+\mathbf{a} \;= \overline{OP} = \mathbf{b}$$
$$[\mathbf{b}+(-\mathbf{a})]+2\mathbf{a} = \mathbf{b}+\mathbf{a} = \overline{OQ}.$$

Both results would perhaps have looked more familiar if we had written $\mathbf{b}+(-\mathbf{a})$ as $\mathbf{b}-\mathbf{a}$. Notice also that $(\mathbf{b}+\mathbf{a})+(\mathbf{b}-\mathbf{a}) = 2\mathbf{b} = \overline{OR}$, where R is the fourth vertex of the parallelogram $OQRQ'$, a reminder of the parallelogram construction so often used for vector quantities in mechanics.

It should be noted that when once we suppose that line segments can be treated in this vectorial way, a whole complex of congruency results (triangle and parallelogram theorems) of Euclidean geometry flows from it; e.g. defining P as the midpoint of QQ' and of OR it follows that $\overline{QR}$ ($= -\mathbf{a}+\mathbf{b}$) is equal and parallel to $\overline{OQ'}$.

EXERCISE 3a (*Geometrical*)

(1)

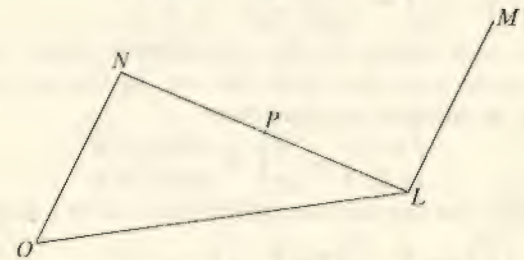


FIG. 9

In the figure, P is the midpoint of LN and LM is equal and parallel to ON in the same sense. Prove that P is the midpoint of OM.

(*Hint*: Choose algebraic symbols for $\overline{OP}$, $\overline{LP}$.)

(2) Writing $\overline{OA} = \mathbf{a}$ and $\overline{OB} = \mathbf{b}$, which of the following forms are correct for $\overline{AB}$?

$$\mathbf{b}-\mathbf{a};\;\; \mathbf{a}+(-\mathbf{b});\;\; -\mathbf{b}+\mathbf{a};\;\; \mathbf{b}+(-\mathbf{a});\;\; -\mathbf{a}+\mathbf{b}$$

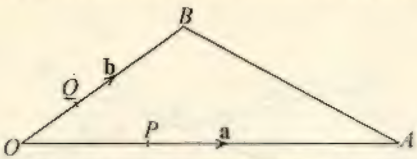


FIG. 10

(3) Given $OP = \frac{1}{3}\,OA$ and $OQ = \frac{1}{3}\,OB$, write $\overline{PQ}$ in terms of **a** and **b** and hence show that $\overline{PQ} \parallel \overline{AB}$. What is the relation of their magnitudes?

(This example indicates how the similarity section of Euclid's geometry flows from the assumptions about vectors with which this chapter began.)

(4) O, A, B, C are points in space not in the same plane. With O as origin the positions of A, B, C are given by the 3-vectors $\overline{OA} = \mathbf{a}$, $\overline{OB} = \mathbf{b}$, $\overline{OC} = \mathbf{c}$. Describe the positions of the points given by $\frac{1}{2}(\mathbf{m}+\mathbf{c})$, $\frac{2}{3}\mathbf{m}+\frac{1}{3}\mathbf{c}$ where $\mathbf{m}=\frac{1}{2}(\mathbf{a}+\mathbf{b})$.

(5) A regular hexagon $ABCDEF$ has centre O. $\overline{OA} = \mathbf{a}$ and $\overline{OC} = \mathbf{c}$. Express in terms of **a** and **c**: (*i*) all the diagonals drawn from A, (*ii*) $\overline{EB}$, $\overline{CF}$, and the join of the midpoint of AF to the midpoint of CD.

3.2 Implicit in the previous work, but necessary to emphasise, is that two vectors cannot be equal unless *all* their separate component numbers are equal; e.g. suppose we are given that $\frac{1}{2}\begin{pmatrix}4\\-6\end{pmatrix} + \begin{pmatrix}p\\5\end{pmatrix} = \begin{pmatrix}0\\q\end{pmatrix}$.

The left-hand side $= \begin{pmatrix}2\\-3\end{pmatrix} + \begin{pmatrix}p\\5\end{pmatrix} = \begin{pmatrix}p+2\\2\end{pmatrix}$.

$$\text{Then } \begin{pmatrix}2+p\\2\end{pmatrix} = \begin{pmatrix}0\\q\end{pmatrix} \Rightarrow \begin{cases}\text{both that } & 2+p = 0\\ \text{and that } & 2 = q.\end{cases}$$

The reader who reflects on this situation will realise that the two statements on the *right* also imply the one on the *left* and we may write $\Leftarrow$ also, or combine the signs thus:

$$\begin{pmatrix}2+p\\2\end{pmatrix} = \begin{pmatrix}0\\q\end{pmatrix} \Leftrightarrow \begin{matrix}2+p = 0\\ \text{and } 2 = q.\end{matrix}$$

We have seen that *two* facts are being told to us by a single equation of the form: $\begin{pmatrix}4\\2\end{pmatrix}x + \begin{pmatrix}3\\-2\end{pmatrix}y = \begin{pmatrix}5\\-8\end{pmatrix}$.

(The multipliers $x+y$ have been written *after* the vectors, but simply for the reason that the next line would otherwise look strange: it is optional whether the multiplier is written fore or aft.)

These two facts, when we have brought them to light, should be enough to determine the values of x and y required to make it a true statement (i.e. to find the solution of the equation). Let us proceed.

$$\begin{pmatrix}4x\\2x\end{pmatrix} + \begin{pmatrix}3y\\-2y\end{pmatrix} = \begin{pmatrix}5\\-8\end{pmatrix}$$

$$\Leftrightarrow \begin{cases} & 4x+3y = 5\\ \textit{and} & 2x-2y = -8\end{cases}$$

$$\Leftrightarrow x = -1 \text{ and } y = 3.$$

$$\textit{Check}:\ -1\begin{pmatrix}4\\2\end{pmatrix} + 3\begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}-4+9\\-2-6\end{pmatrix} = \begin{pmatrix}5\\-8\end{pmatrix}.$$

Let us illustrate the check on a diagram, naming the given vectors **a**, **b** and **c**; e.g. $\mathbf{a} = \begin{pmatrix}4\\2\end{pmatrix}$.

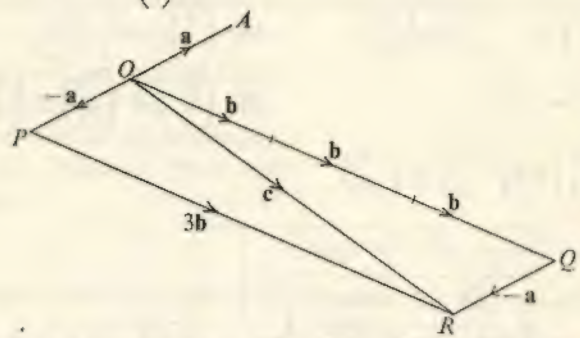


FIG. 11

The result is shown by either $\overline{OP}+\overline{PR} = \overline{OR}$ or $\overline{OQ}+\overline{QR} = \overline{OR}$.

It is important to notice a further point, viz. we could have used a geometrical construction to *get* the multipliers (-1) and 3, not merely to check them.

Thus:

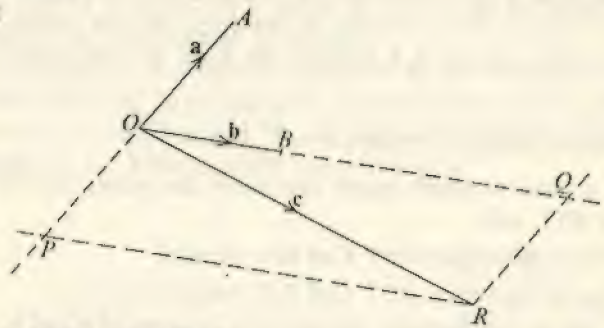


FIG. 12

Draw **a**, **b** and **c** as shown.

Through R draw parallels to **a** and **b**, to meet the lines OA and OB, produced as necessary, at P and Q, and complete the parallelogram. Then *measure* off OA into OP (once backwards) and OB into OQ (3 times).

EXERCISE **3b** (*Logical and algebraic*)

(1) Find three other examples in which a single statement implies two statements. One might be:

Edward and Mary are the king and queen.

$$\Rightarrow \begin{cases} \text{Edward is the king.} \\ \text{Mary is the queen.} \end{cases}$$

This example is reversible: which of yours are also?

(2) In each of the following vector equations, find, where possible, the values to be given to x, y.

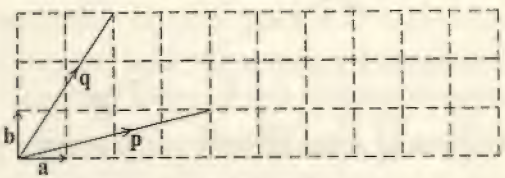
(*i*) $\begin{pmatrix} 2x \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2y \end{pmatrix}$ (*ii*) $\begin{pmatrix} 3 \\ x \end{pmatrix} = y\begin{pmatrix} x \\ 12 \end{pmatrix}$

(*iii*) $\begin{pmatrix} 2 \\ x \end{pmatrix} = y\begin{pmatrix} 3 \\ x \end{pmatrix}$ (*iv*) $\begin{pmatrix} 2x \\ 1 \end{pmatrix} + \begin{pmatrix} 3y \\ 2y \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$

(*v*) $x\begin{pmatrix} x \\ 3 \end{pmatrix} + 2\begin{pmatrix} -1 \\ x \end{pmatrix} = \begin{pmatrix} x \\ 10 \end{pmatrix}$

(3)

FIG. 13



Copy the above diagram on graph paper.

Mark on the paper two distinct directed lines which represent $\mathbf{p}-\mathbf{q}$, one which gives $\mathbf{q}-\mathbf{p}$, one for $\mathbf{p}+\frac{1}{2}\mathbf{q}$, and one for $\frac{1}{2}(\mathbf{p}-\mathbf{q})$.

Express all the above vectors in the form $\lambda\mathbf{a}+\mu\mathbf{b}$ where λ, μ are numbers. Also express $\mathbf{a}$, $\mathbf{b}$ each in the form $l\mathbf{p}+m\mathbf{q}$ where l, m are numbers.

Check your results by writing $\mathbf{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

(4) Take a piece of graph paper and mark the points $A(1, 0)$ $B(3, 1)$ $C(5, -2)$ $D(3, -3)$.

(*a*) What is the vector $\overline{AB}$? Call this vector **a**.

(*b*) What is the vector $\overline{BC}$? Call this vector **b**.

To the rest of this question give answers in terms of **a** and **b**, and not in component form:

(*c*) What vectors are $\overline{AD}$, $\overline{BC}$, $\overline{DC}$, $\overline{BA}$, $\overline{CB}$, $\overline{AC}$, $\overline{CA}$?

If E is the point of intersection of the diagonals of the quadrilateral, what are $\overline{AE}$, $\overline{EC}$, $\overline{BE}$, $\overline{ED}$?

3.3 *Worked examples from mechanics*

(*a*) When a man travels at 10 m.p.h. due north, he notices that the wind appears to come from the east with a speed of 10 m.p.h. Find the true speed and direction of the wind.

An appeal to books on mechanics will tell us that if $\mathbf{v}$ is the velocity of the man relative to the ground, and $\mathbf{V}$ is the velocity of the wind relative to the ground, then the velocity of the wind as observed by the man is $\mathbf{V}-\mathbf{v}$.

In this question we know $\mathbf{v}$ and $(\mathbf{V}-\mathbf{v})$, and we have to find $\mathbf{V}$; clearly we have only to add the two vectors which we know:

$$\mathbf{V} = (\mathbf{V}-\mathbf{v})+\mathbf{v}.$$

Putting numbers into the figure we immediately see that the velocity of the wind is approximately 14·1 m.p.h. blowing from the south-east.

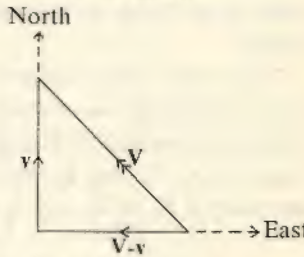


FIG. 14

We might also have solved this question by taking axes east and north so that $\mathbf{v} = \begin{pmatrix} 0 \\ \cdot 10 \end{pmatrix}$, $\mathbf{V}-\mathbf{v} = \begin{pmatrix} -10 \\ 0 \end{pmatrix}$.

Hence $\mathbf{v}+(\mathbf{V}-\mathbf{v}) = \begin{pmatrix} -10 \\ 10 \end{pmatrix} = \mathbf{V}$ and the result follows quickly once again. However it is not usually an advantage to use vector components with this type of problem—it is better to use the sine or cosine formula in the vector diagram formed as shown.

(*b*) A weight of 10 lb weight is supported by two strings, one inclined at 45° and one at 60° to the vertical. Find the sizes of the tensions. In this instance the three vectors, $\mathbf{T}_1$, $\mathbf{T}_2$ and 10 lbf downwards, must be in equilibrium, i.e. have resultant (sum-vector) **0**.

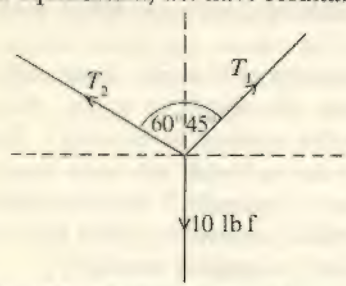


FIG. 15 Diagram showing situation

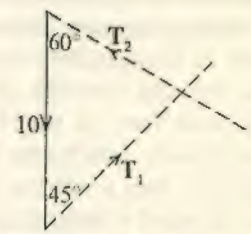


FIG. 16 Vector Diagram

The directions of $\mathbf{T}_1$ and $\mathbf{T}_2$ are defined in the situation diagram, and the vectors themselves (with sum zero) are shown in the second diagram. The fact that the sum is zero is of course the well-known triangle of forces result.

We now use the sine formula $\dfrac{10}{\sin 75^\circ} = \dfrac{T_2}{\sin 45^\circ} = \dfrac{T_1}{\sin 60^\circ}$

$$\Leftrightarrow \quad T_2 = 7{\cdot}16 \text{ lbf}$$
$$\text{and } T_1 = 8{\cdot}76 \text{ lbf.}$$

It is important to note that a vector can only represent the magnitude and direction of a force: the force itself has also a point of application. (The three forces in the above example were all applied at the same point.)

The reader who studies mechanics by vector methods will find each force defined by *two* vectors, (*i*) a position vector $\mathbf{r}$ to specify the point of application and (*ii*) $\mathbf{F}$ to give its magnitude and direction. Two forces given by $\mathbf{r}_1, \mathbf{F}_1$ and $\mathbf{r}_2, \mathbf{F}_2$ are equivalent (in rigid-body mechanics) if and only if $\mathbf{F}_1 = \mathbf{F}_2$ and $\mathbf{r}_1 - \mathbf{r}_2$ is parallel to $\mathbf{F}_1$ and $\mathbf{F}_2$.

EXERCISE 3c
(including mechanics and other applications)

(1) A ship steaming at 30 knots (i.e. 30 nautical miles per hour) on a course of 060° reports a submarine to be due north of it, remaining due north and closing its distance at a rate of 10 knots.

Express the following velocities as 2-vectors using easterly and northerly components:

(*i*) the velocity $\mathbf{u}$ of the ship; (*ii*) the velocity $\mathbf{v}$ of the submarine as seen from the ship as origin; (*iii*) the true velocity of the submarine. Give also an expression for the third in terms of $\mathbf{u}$ and $\mathbf{v}$.

Supposing the submarine to be rising towards the surface at a rate of 1 knot, state all the velocities suitably as 3-vectors. What is the actual speed of the submarine?

(2) A man is cycling due north with a velocity of 10 m.p.h., into a wind which is blowing from the north with a velocity of 15 m.p.h. With what velocity does the cyclist observe the wind to be blowing? The wind now veers so that it blows from the east with the same speed: what velocity does it now have according to the cyclist? Between which directions may a wind of this strength blow so that the cyclist can say that the wind is behind him? (See worked example if necessary.)

(3) A distant star is in the plane of the orbit of the earth (E) about the

sun (S). The velocity $\mathbf{v}$ of E is 18 miles per second in a direction perpendicular to ES, and the light of the star approaches E with true velocity $\mathbf{V}$ equal to 180,000 miles per second exactly in direction ES.

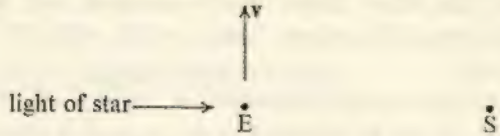


FIG. 17

Show that the apparent direction of the light will differ from the true direction by about $\frac{1}{3}$ of a minute of arc and show on a diagram which way the error will be.

(4) The figure shows, in plan, a 120-ft mast with foot at M and cables fastened to points A, B, C on the ground. The cables, which may be treated as straight, run to the top T of the mast.

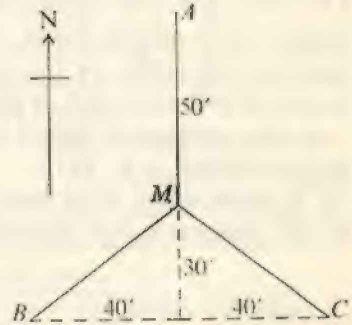


FIG. 18

Calculate (*i*) the lengths of the cables and (*ii*) the 3-vectors $\overline{TA}$, $\overline{TB}$, $\overline{TC}$ referred to west, north and *down*ward as axes.

(*iii*) State also as vectors the tensions $\mathbf{P}, \mathbf{Q}, \mathbf{R}$ in the cables TA, TB, TC, given that the vertical component of $\mathbf{P}$ is 1200 lb and that the vector sum $\mathbf{P}+\mathbf{Q}+\mathbf{R}$ is vertical (as would be so in still air).

(*Hints*: (*a*) If two vectors $\mathbf{a}, \mathbf{b}$ have the same direction then $\mathbf{a} = k\mathbf{b}$. (*b*) The equality of two 3-vectors implies three numerical equations.)

(5) The make-up of every food can be expressed as a vector, the first three components being grams of edible protein, fat and carbohydrate, and the fourth the calories of energy produced, in every case for 500 g of the food. Typical vectors (written as rows for convenience) are:

Meat	$\mathbf{m} = (100, 140, 0, 1900)$
Bread	$\mathbf{b} = (30, 0, 240, 1230)$
Eggs	$\mathbf{e} = (60, 50, 6, 830)$
Cheese	$\mathbf{c} = (140, 180, 12, 2550)$
Butter	$\mathbf{t} = (0, 450, 0, 4500)$
Sugar	$\mathbf{s} = (0, 0, 500, 2250)$

and a man's daily requirement is $\mathbf{r} = (90, 90, 480, 3500)$.

Suggest a basic diet for a man if he eats no cheese, but exactly 50 g butter and 250 g of sugar per day. He will not eat more than 500 g of bread, or more than 250 g of meat.

We suppose that other dietary needs—minerals and vitamins—are met. (*Hint*: Disregard the calorie figures, which will tally if the others do. Suggest why.)

(6) In mental testing, a pupil's results in a 'battery' of 3 tests is called his 'test profile'. It can be written $\mathbf{c} = (c_1, c_2, c_3)$; it may be regarded as a vector since it is claimed that it is the sum of two parts **a**, **b**, these being the portions of his score assignable to his general ability and to his special (verbal) ability respectively. For a revised battery his profile **d** is said to give a 50% greater score to his general and 50% less (than before) to his special ability. Express **d** in terms of **a** and **b**. What would be his profile on the *mean* of the two batteries (test 1 being combined with test 1 and so on)?

In what proportion should **c** and **d** be combined, if the result is to be proportional to $\mathbf{a}+\frac{1}{2}\mathbf{b}$?

(7) A recent model of car has a rear window which slopes at an angle of 100° backwards from the direction of travel.

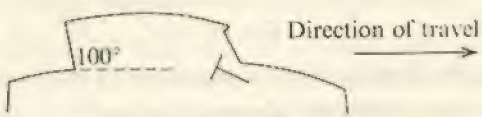


FIG. 19

When the car is travelling at 30 m.p.h. and rain is falling at 30° to the vertical, what is the smallest velocity that the rain may have in order to wet the back window?

chapter 4

VECTORS AND LINEAR EQUATIONS

4.1 We have mentioned that computers are often concerned with vectors of great size. These may, for example, be stock-holding vectors, giving the numbers in stock of each of a standard series of items:

$$(253 \quad 140 \quad 31 \quad 87 \ldots)$$

There is no doubt that this satisfies our definition: if a new order, cast in the form of a similar vector, is delivered to this depot, the sum (formed in the defined way) represents the new stock.

In technology, however, the large vectors which occur are more often made up of the coefficients of equations. We have considered a simple case of this in exercise 3c, question 5, and we shall now carry the matter further, remembering that we normally expect a unique solution, i.e. a single set of solutions for the unknowns x, y, etc.

We shall consider first a very simple case—two linear equations in two unknowns, i.e.

$$\begin{aligned} a_1x+b_1y &= c_1 \\ a_2x+b_2y &= c_2 \end{aligned} \qquad \text{involving six coefficients.}$$

The coefficients of x form our first vector, $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ which we write shortly as **a**: similarly $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathbf{b}$ and $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{c}$. Then we can restate our given equations:

$$\left.\begin{aligned} a_1x+b_1y &= c_1 \\ a_2x+b_2y &= c_2 \end{aligned}\right\} \Leftrightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} x + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} y = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$
$$\Leftrightarrow \mathbf{a}x+\mathbf{b}y = \mathbf{c}.$$

We shall proceed to solve a numerical example.

4.2 Solve the equation $\mathbf{a}x+\mathbf{b}y = \mathbf{c}$,

where $\mathbf{a} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -6 \\ 2 \end{pmatrix}$ and $\mathbf{c} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}$.

In full for an elementary algebraic solution, we have:

$$\left.\begin{aligned}3x-6y &= 10\\ -x+2y &= 4\end{aligned}\right\}.$$

To eliminate x, add three times the second equation to the first. The reader is warned against solving such equations by substitution from one such equation into the other. It is not incorrect, of course, but in all but the very simplest cases it is clumsy and very liable to error. In practical examples with non-integral coefficients it is *never* done.

We find y disappears as well, and the result is $0 = 22$.

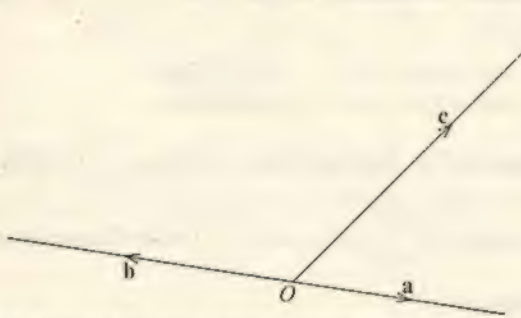
Condensing logically what we have found:

$$\left.\begin{aligned}3x-6y &= 10\\ -x+2y &= 4\end{aligned}\right\} \Rightarrow 0 = 22.$$

There is nothing wrong with the argument. If x and y have values which satisfy the two given statements simultaneously, then $0 = 22$. Clearly, therefore, there are no such values: we can find values which satisfy *either* but not both. The statements, i.e. the equations, are said to be INCONSISTENT. (This is clearly seen if the loci are drawn as in co-ordinate geometry.)

If we attempt to carry out the geometrical construction described in section 3.2 to solve the equation $\mathbf{a}x+\mathbf{b}y = \mathbf{c}$, the reason for failure is apparent:

FIG. 20



a and **b** lie along the same 'direction'. (We have shown them in the same line, because we chose to draw all the vectors from the same origin O.) Any combination of **a**'s and **b**'s will accordingly lie in that 'direction' and cannot, in general, be equal to **c**.†

In section 3.2 the solution by construction depended on our being able to draw a parallelogram, and this has collapsed.

† If **c** also is in that direction or is zero we *can* do it, but a little thought shows that this is too easy: **c** can then be expressed in an infinite variety of ways. We fail to get a *unique* solution, but the equations are now CONSISTENT.

This simple example leads us to make a tentative statement for 2-vectors in general, viz.:

$$\left.\begin{array}{l}\text{Equation } \mathbf{a}x+\mathbf{b}y = \mathbf{c} \text{ has a}\\ \text{unique solution in } x \text{ and } y\end{array}\right\} \Leftrightarrow \left\{\begin{array}{l}\mathbf{a} \text{ and } \mathbf{b} \text{ not in the same}\\ \text{or opposite directions.}\end{array}\right.$$

But the right-hand statement is unsatisfactory. It is not in accord with our policy of using algebraic descriptions to attain precision. This is not a fad: has every 2-vector got a direction? What about **0**? (The student should try for himself to restate the right-hand side before reading on.)

Possible statements to try out are:

(*i*) **b** cannot be expressed as a numerical multiple of **a**.
a cannot be expressed as a numerical multiple of **b**.

(*ii*) If $\mathbf{a} = \begin{pmatrix}a_1\\ a_2\end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix}b_1\\ b_2\end{pmatrix}$ then $\dfrac{a_1}{a_2} \neq \dfrac{b_1}{b_2}$.

The statement can be tidied up by following the lines of either (*i*) or (*ii*). We shall develop both lines of approach, by total vectors and by components, in the next section.

4.3 (*a*) *Linear independence of vectors.* This develops approach (*i*) of the previous section.

If numbers λ, μ exist (not both zero) such that $\lambda\mathbf{a}+\mu\mathbf{b} = \mathbf{0}$, then the vectors are said to be *linearly dependent*. Otherwise, they are *linearly independent*.

Then the relationship we require would seem to be:

$$\left.\begin{array}{l}\mathbf{a}x+\mathbf{b}y = \mathbf{c} \text{ has a unique}\\ \text{solution in } x, y\end{array}\right\} \Leftrightarrow \left\{\begin{array}{l}\mathbf{a}, \mathbf{b} \text{ are linearly}\\ \text{independent}\end{array}\right.$$

where **c** is a non-vector zero.

There is no difficulty in seeing, for the case of two equations, that this is fully satisfactory. It is helpful to remember the result in this form: that if we succeed in getting (non-trivial) λ, μ to fit $\lambda\mathbf{a}+\mu\mathbf{b} = \mathbf{0}$ then we fail in getting (uniquely) x, y, to satisfy $x\mathbf{a}+y\mathbf{b} = \mathbf{c}$. The term *non-trivial* is a useful one, to exclude the 'trivial' solution $\lambda = 0, \mu = 0$.

(*b*) *An expression in terms of components, which tests our equations*

We cannot use $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2}$ as a general test, because this is meaningless when a_2 is zero. (In fact, we *like* our vectors to be parallel to the axis of our 'graph-paper' for convenience, so $a_2 = 0$ is quite a common occurrence.)

The way out is to write $a_1b_2-a_2b_1=0$. (The student should verify that this secures, in every case where the term has meaning, that the vectors are parallel.) This is an important expression, and as it is derived from the pattern $\begin{matrix} a_1 & b_1 \\ a_2 & b_2 \end{matrix}$ of the coefficients, it is commonly written in square form between ruled vertical lines, thus:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2-a_2b_1.$$

Such an expression is called a *determinant of order 2*. Then we can write the two-way implication:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}x+\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}y=\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \Leftrightarrow \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$$

has a unique solution in x, y.

Note that in our example of section 3.2 the determinant was

Equation $\begin{vmatrix} 4 & 3 \\ 2 & -2 \end{vmatrix} = -8-6=-14,$

and in 4.2 $\begin{vmatrix} 3 & -6 \\ -1 & 2 \end{vmatrix} = 6-(-1)(-6)=6-6=0.$

(*Note*: The reader will already be familiar with an example of an algebraic expression which by its value gives information about the character of an equation: the quadratic† equation $ax^2+bx+c=0$ is said to have the *discriminant* b^2-4ac. If this is zero the equation has equal roots; if positive, real roots; if a perfect square, rational roots; if negative, complex roots.)

In our example, the determinant, by being zero or not, discriminates, but its sign does not matter: the equations in 3(*c*) could have come in the other order, giving: $\begin{vmatrix} 3 & 4 \\ -2 & 2 \end{vmatrix} = 6+8=+14.$

EXERCISE 4a

(1) Evaluate the determinants:

(*a*) $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}$ (*b*) $\begin{vmatrix} 8 & 2 \\ -3 & 6 \end{vmatrix}$ (*c*) $\begin{vmatrix} 4 & 2 \\ -1 & 8 \end{vmatrix}$ (*d*) $\begin{vmatrix} 3 & 4 \\ -6 & -8 \end{vmatrix}$.

† We have to know that it really is a quadratic, i.e. $a \neq 0$. In computer work in which a supposedly quadratic equation appears at some stage, a test has to be applied to ensure that $a \neq 0$ before carrying out the discriminant procedure.

(2) Find, by geometrical construction on graph paper, multipliers p, q, such that $p\begin{pmatrix} 2 \\ 1 \end{pmatrix}+q\begin{pmatrix} 3 \\ 4 \end{pmatrix}=\begin{pmatrix} 4\frac{1}{2} \\ 3\frac{1}{2} \end{pmatrix}$.

(3) Evaluate the determinants:

(*a*) $\begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix}$ (*b*) $\begin{vmatrix} 8 & -3 \\ 2 & 6 \end{vmatrix}$ (*c*) $\begin{vmatrix} 4 & -1 \\ 2 & 8 \end{vmatrix}$ (*d*) $\begin{vmatrix} 3 & -6 \\ 4 & -8 \end{vmatrix}$.

(4) Which of the following pairs of vectors are linearly dependent? For those which are so, find values of λ and μ (not both zero) such that $\lambda\mathbf{a}+\mu\mathbf{b}=\mathbf{0}$. (Note that if λ, μ is a solution, so also is 2λ, 2μ and so on.)

(*i*) $\mathbf{a}=\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\mathbf{b}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ (*ii*) $\mathbf{a}=\begin{pmatrix} 4 \\ -3 \end{pmatrix}$, $\mathbf{b}=\begin{pmatrix} -8 \\ 6 \end{pmatrix}$

(*iii*) $\mathbf{a}=\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{b}=\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (*iv*) $\mathbf{a}=\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\mathbf{b}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(5) Show directly, and also by an argument on dependence of vectors, that $\begin{vmatrix} m^2 & n^2 \\ m & n \end{vmatrix}$ has $m-n$ as a factor.

(6) Prove that, if $\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ are linearly independent, then so are the vectors $\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$ and $\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$.

(7) Evaluate the determinants $\Delta_1=\begin{vmatrix} 3 & 4 \\ 1 & -1 \end{vmatrix}$ and $\Delta_2=\begin{vmatrix} 3k & 4k \\ 1 & -1 \end{vmatrix}$.

What conclusion do you draw about the presence of k? Use your result to simplify the evaluation of

(*a*) $\begin{vmatrix} 38 & 133 \\ 2 & 7 \end{vmatrix}$ (*b*) $\begin{vmatrix} 16 & 64 \\ 18 & 90 \end{vmatrix}$ (*c*) $\begin{vmatrix} 21 & 28 \\ 26 & 39 \end{vmatrix}$.

(8) Decide which of the following pairs of equations have unique solutions, and say which of the others are consistent or inconsistent.

(*a*) $3x-2y=5$, $2x-y=4$ (*b*) $3x-4y=7$, $-6x+8y=-14$

(*c*) $2x+16y=7$, $-x-8y=2$ (*d*) $5x+3y=6$, $10x+6y=12$

(*e*) $2x+3x=5$, $4x+6x=10$ (*f*) $x+2y=4$, $2x+5y=10$

(9) What value of λ makes the vectors $\begin{pmatrix}1\\3\end{pmatrix}\begin{pmatrix}-2\\\lambda\end{pmatrix}$ linearly dependent?

What value must μ take to make the equation $x\begin{pmatrix}1\\3\end{pmatrix}+y\begin{pmatrix}-2\\\lambda\end{pmatrix}=\begin{pmatrix}5\\\mu\end{pmatrix}$ consistent?

(10) Find all the values of λ which cause the equations $\begin{matrix}\lambda x+2y = 5\\ x+2\lambda y = \mu\end{matrix}$ not to have unique solutions, and determine μ for consistency in each case.

(11) (Referring to section 4.3(*a*))

Given three 2-vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, where $\mathbf{c} \neq \mathbf{0}$, let us assume both the following statements to be true:

(*i*) For some λ, μ (not both zero) $\lambda\mathbf{a}+\mu\mathbf{b} = \mathbf{0}$.

(*ii*) There exist unique x, y such that $x\mathbf{a}+y\mathbf{b} = \mathbf{c}$.

The reader should obtain directly from these statements a deduction which is inconsistent with the assumptions.

(12) (An exercise in logic based on question 11)

Giving the symbol d to statement (*i*), affirming linear dependence of $\mathbf{a}$, $\mathbf{b}$; and the symbol s to statement (*ii*), affirming (unique) solubility of the equation, deduce which of the following are valid implications:

$$d \Rightarrow s; \quad d \Rightarrow \text{not-}s; \quad s \Rightarrow d; \quad s \Rightarrow \text{not-}d.$$

(13) If p and q are any two statements, then by definition the assertion '$p \Rightarrow q$' is equivalent to 'p and not-q cannot both be true'. Show that it is also equivalent to 'not-$q \Rightarrow$ not-p', and illustrate by an example.

4.4 *The layout of the solution (for integral coefficients)*

For all but the very simplest situations, such as could almost be solved and checked mentally, it is vital to adopt a systematic layout of the solution.

First of all, since the arithmetical work is carried out on the coefficients, we require to display only these and not their attendant x's or y's. By the use of columns and guide-lines we can see which coefficient is which: e.g. the equations $\begin{cases}x-\ y = -4\\ 4x+3y = \ \ 5\end{cases}$ would in fact be written as

$$\begin{array}{rr|r} 1 & -1 & -4 \\ 4 & 3 & 5 \\ \hline \end{array}.$$

Secondly, at each stage of the solution we have a new pair of equations with the same solutions (if any) as before. In the above case we might form a new equation by multiplying the second by 2 and subtracting it from the first. The new equation is written in place of one of the former (along with an indication of how it was obtained). The pattern will therefore look thus, with the new pair separated from the old by a horizontal line:

$$\begin{array}{l rr|r l} & 1 & -1 & -4 & \\ & 4 & 3 & 5 & \\ \cline{2-4} r_1' = r_1 & 1 & -1 & -4 & \\ r_2' = r_2-4r_1 & 0 & 7 & 21 & \Rightarrow\ y = 3 \text{ etc.} \end{array}$$

The abbreviations on the left indicate the steps by which the new rows (primed) are derived from the old, e.g. $r_2' = r_2-4r_1$ means 'new-row-2 = old-row-2 minus four times old-row-1'. (We commonly omit the repetition statement $r_1' = r_1$: the important thing is that the equation should be repeated.)

At each stage of the solution we have a new set of equations which supersedes the old (and in computer practice the new coefficients actually take the places formerly occupied by the old, in order to economise storage space). It is as if we had now been asked to solve the equations

$$\begin{aligned} x-y &= -4\\ 7y &= 21. \end{aligned}$$

In this situation we have no eliminating to do: the value of y leads to the value $x = -1$ by back-substitution. In the general case we have $\left.\begin{aligned}a_1x+b_1y &= c_1\\ b_2y &= c_2\end{aligned}\right\}$ in which we assume c_1, c_2 not both zero.

We have a solution for y provided $b_2 \neq 0$, and proceed to a value of x, for which we require $a_1 \neq 0$. In fact we must have $a_1b_2 \neq 0$: which, for this set of equations, is the *determinant of the left-hand-side coefficients* (since $a_2b_1 = 0$).

The process by which the complete array of coefficients is transformed into another array is called a *row-operation*. Only one row-operation was required above, which left one row unchanged and produced a zero at the left of the other. We shall study such operations in detail in the next chapter, where it will be seen that the aim is to replace certain of the coefficients by zeros, working from the left-hand corner, until the

remainder of the solution is entirely back-substitution: for example, we might have:

$$\left.\begin{array}{ccc|c} 2 & 1 & 2 & 0 \\ 0 & 3 & 1 & 9 \\ 0 & 0 & 2 & 6 \end{array}\right. \Leftrightarrow \begin{cases} z = 3 \\ y = 2 & \text{(using equation II)} \\ x = -4 & \text{(from equation I).} \end{cases}$$

Such an array of coefficients, with zeros below the main diagonal, is called an *echelon form.*

The reader should note that it may be convenient, in order to achieve the result smoothly, to interchange the equations at the outset so as to get a suitable coefficient of x at the top—as was already the case in the example ($a_1 = 1$ means that row 1 can certainly be multiplied so as to eliminate the x-coefficient from the other equation).

EXERCISE 4b

Solve the following sets of equations, using the layout described:

(1) $x+5y = 18$
$3x+4y = -1$

(2) $\begin{array}{cc|c} 2 & -7 & 4 \\ 8 & -5 & 62 \end{array}$

(3) $\begin{array}{cc|c} 6 & 7 & 29 \\ -1 & 3 & 16 \end{array}$

(4) $\begin{array}{cc|c} 7 & 9 & 1 \\ 3 & 5 & -3 \end{array}$

(*Hint* for question 4: Use $r_1' = r_1 - 2r_2$, $r_2' = r_2$ to get a simple x-coefficient—an easier start, although it involves an extra stage in the solution.)

4.5 *The layout for non-integral coefficients*

For this case, which is so important in practice, we have the same general pattern with two new features:

(*i*) We put at the top the equation with the *numerically-largest* coefficient of x. (This is called the *pivotal* equation.)

(*ii*) We have a further column in the layout which serves to detect any error as soon as possible after it occurs. In this column we place, for each equation, the sum of all the coefficients (including the right-hand-side coefficient). The row-operation is carried out also on this extra number: how the check operates will be seen in the following example.

Worked example

To solve $\begin{cases} 0{\cdot}50x+0{\cdot}32y = 1{\cdot}78 \\ 0{\cdot}10x+0{\cdot}70y = 2{\cdot}31 \end{cases}$, treating the given figures as exact.

The pivotal equation is already at the top, and the array is:

				Sum-column	
	0·50	0·32	1·78	2·60	
	0·10	0·70	2·31	3·11	
	0·50	0·32	1·78	2·60	
$r_2' = r_2 - 0{\cdot}2r_1$	0	0·636	1·954	2·590	(checks $\surd$)

(*Note*: 2·590 is derived from $3{\cdot}11 - 0{\cdot}2 \times 2{\cdot}60$, but it is also found to be the sum of the numbers in its row.)

Then $y = 1{\cdot}954 / 0{\cdot}636 = 3{\cdot}072$ to 4 figures and $0{\cdot}50x = 1{\cdot}78 - 0{\cdot}32 \times 3{\cdot}072 = 0{\cdot}8072$, giving $x = 1{\cdot}614$ to 4 figures.

The results check to within 0·001 in the original equations. (The question of the reliability of these results if the data are only approximate will not be considered here—it belongs to numerical analysis; so also does the reason for the choice of the pivotal equation, which has the effect of giving a multiplier, 0·2 in the above case, of magnitude less than unity.)

The above procedure is ideally suited for work with a desk machine; but it is also profitable to carry out some examples with a slide-rule (with an intermediate line of figures written in the layout, $0{\cdot}2r$ in the above case).

EXERCISE 4c

Solve the following equations approximately, with the aid of a slide-rule, checking your results:

(1) $1{\cdot}08x+2{\cdot}56y = 1{\cdot}96$
$-0{\cdot}49x+1{\cdot}82y = 0{\cdot}97$

(2) $2{\cdot}15x-3{\cdot}20y = 2{\cdot}26$
$1{\cdot}23x+0{\cdot}80y = 1{\cdot}03$

(3) Recalculate the above, using either 4-figure logarithms or, preferably, a desk-machine. In the former case the discrepancy on substitution should not exceed 0·002.

chapter 5

VECTORS AND LINEAR EQUATIONS (3-VECTORS)

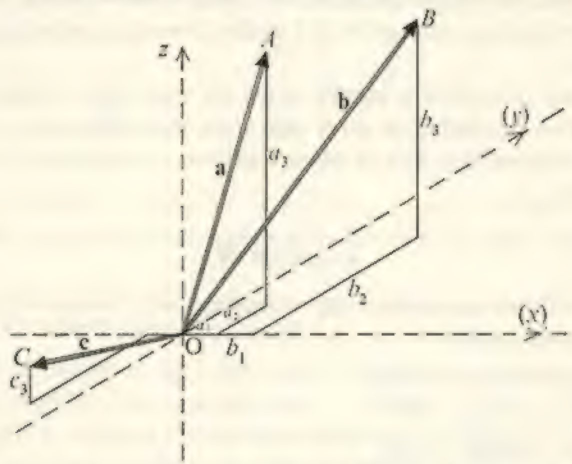
5.1 We shall in this chapter extend our discussion to 3 equations in 3 unknowns, starting as before with a particular case in order to get the feel of the problem:

$$\begin{aligned} x+2y-\ z &= 0 \\ 2x+7y-5z &= 2 \\ 7x+6y+\ z &= 1 \end{aligned} \quad \text{which can be condensed as} \quad \begin{pmatrix}1\\2\\7\end{pmatrix}x+\begin{pmatrix}2\\7\\6\end{pmatrix}y+\begin{pmatrix}-1\\-5\\1\end{pmatrix}z=\begin{pmatrix}0\\2\\1\end{pmatrix}$$

or $\mathbf{a}x+\mathbf{b}y+\mathbf{c}z = \mathbf{d}$.

To consider this geometrically we have to use 3 dimensions. The drawing attempts to represent this, but the student would be well advised to try it out in space.

FIG. 21



It will be seen that any vector *in the plane AOB* can be obtained by adding vectors $\lambda\mathbf{a}$ and $\mu\mathbf{b}$ (since $\mathbf{a}$, $\mathbf{b}$ are not in the same line) where λ, μ have suitably chosen values. In order, with the help of $\nu\mathbf{c}$, to make the final result equal to the vector $\mathbf{d}$, it is essential that $\mathbf{c}$ itself should not also be in the plane AOB. (If it were so, then any combination of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ would also be so.)

The feel of this situation can be obtained by considering it as a force problem, in which a resultant force $\mathbf{d}$ (at O) is to be obtained by adjusting the magnitudes of forces in the directions of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ at O. It is clear that $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ must not be coplanar in order that the adjustment can be made (except in the special cases $\mathbf{d}$ coplanar with $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$; or $\mathbf{d} = \mathbf{0}$).

An experimental or intuitive approach leads us to suspect that the following situation exists, provided $\mathbf{d} \neq \mathbf{0}$, viz:

$\mathbf{a}x+\mathbf{b}y+\mathbf{c}z = \mathbf{d}$ is uniquely soluble for x, y, z

$\Leftrightarrow$ $\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{c} = \mathbf{0}$ is NOT satisfied by any set of values of λ, μ, ν other than the trivial 0, 0, 0; (and the latter statement can by definition be written as:

$\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *are linearly independent*).

EXERCISE 5a

(1) Show (by actually finding numbers λ, μ, ν such that $\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{c} = 0$) that the following vectors are linearly dependent, and sketch the vectors in a diagram in each case:

(*i*) $\mathbf{a} = \begin{pmatrix}1\\2\\0\end{pmatrix}$, $\mathbf{b} = \begin{pmatrix}2\\0\\4\end{pmatrix}$, $\mathbf{c} = \begin{pmatrix}0\\1\\-1\end{pmatrix}$;

(*ii*) $\mathbf{a} = \begin{pmatrix}2\\-1\\0\end{pmatrix}$, $\mathbf{b} = \begin{pmatrix}3\\4\\1\end{pmatrix}$, $\mathbf{c} = \begin{pmatrix}-6\\3\\0\end{pmatrix}$.

(2) Show that $\begin{pmatrix}2\\-1\\0\end{pmatrix}$, $\begin{pmatrix}3\\4\\0\end{pmatrix}$, $\begin{pmatrix}10\\0\\1\end{pmatrix}$ are linearly independent.

(*Hint*: Try to find λ, μ, ν and show that the required equations are insoluble except for $\lambda = \mu = \nu = 0$.)

(3) Verify the connection of linear independence with unique solubility, using the vectors of question 2 above and solving $\mathbf{a}x+\mathbf{b}y+\mathbf{c}z = \mathbf{d}$,

where $\mathbf{d} = \begin{pmatrix}44\\0\\5\frac{1}{2}\end{pmatrix}$.

(4) Verify the connection between linear dependence and failure (of one sort or another) to get unique solutions, using the vectors of question (*i*), and inventing vector **d** in the following way:

Choose simple values of x, y, z to be solutions of your equations, and proceed to calculate $d_1 = a_1x+b_1y+c_1z$; $d_2 = a_2x+b_2y+c_2z$.

If you calculate d_3 in the same way, then you have made the equations CONSISTENT: any failure must be in uniqueness. If, however, you choose d_3 differently you have INCONSISTENCY.

(5) Construct examples of linearly *dependent* 3-vectors, e.g. by the following procedure: write two vectors **a**, **b** which are independent (i.e. $p\mathbf{a} \neq q\mathbf{b}$) and form a vector $\mathbf{c} = \lambda\mathbf{a}+\mu\mathbf{b}$. Show that in general the alteration of *one* of the components of your **c**-vector has the effect of making **a**, **b**, **c** into an *independent* set. (At this point the student will appreciate the need for developing simple tests to discover whether a given set of vectors is linearly dependent or not.)

5.2 We should like to be able to write down, as for the 2-vector case, a determinant which (by being zero or not) would quickly give us the information that we want about the equations. It must, of course (*i*) involve all the coefficient-values combined in some way, and (*ii*) recognise that the place of each coefficient in the total 3-by-3 pattern matters.†

We might in fact decide to write the determinant, usually symbolised by Δ, thus:

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & 7 & -5 \\ 7 & 6 & 1 \end{vmatrix} \quad \text{or in general } \Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

but this is no help until we know a rule for working it out! The reader may have had the curiosity to try to solve the equations with which we began. If so, he has failed: this is a case in which **a**, **b** and **c** are linearly dependent and a carefully-made model of the geometrical figure shows $OABC$ to be a plane (and OD is oblique to it).

We shall not immediately give the rule for finding the value of Δ, but concentrate on the processing of the equations using row-operations just as we did for the simpler case of two equations (in chapter 4).

† Such a pattern is called a 3×3 matrix.

$$\begin{array}{rrr|r} 1 & 2 & -1 & -1 \\ 2 & 7 & -5 & 3 \\ 7 & 6 & 1 & -6 \\ \hline 1 & 2 & -1 & -1 \\ 2 & 7 & -5 & 3 \\ 0 & -8 & 8 & 1 \\ \hline 1 & 2 & -1 & -1 \\ 0 & 3 & -3 & 5 \\ 0 & -8 & 8 & 1 \\ \hline 1 & 2 & -1 & -1 \\ 0 & 3 & -3 & 5 \\ 0 & 0 & 0 & 14\frac{1}{3} \\ \hline \end{array} \qquad \begin{array}{l} \\ \\ \\ \\ \\ r_3' = r_3-7r_1 \\ \\ r_2' = r_2-2r_1 \\ \\ \\ \\ r_3' = r_3+\frac{8}{3}r_2 \end{array}$$

It is immediately apparent that no solution is possible, because the last equation is $0.z = 14\frac{1}{3}$, whereas $0.z = 0$ for any value we might suggest for z.

It can be seen that there would be the same kind of trouble if any of the other numbers in what is called the main diagonal were to vanish:

$$\begin{pmatrix} a_1 & \cdot & \cdot \\ 0 & b_2 & \cdot \\ 0 & 0 & c_3 \end{pmatrix}$$

Our condition is in fact $a_1b_2c_3 \neq 0$, i.e. when the matrix of coefficients has been worked into echelon form the product of the elements in the main diagonal determines the nature of the set of equations. We shall later be defining an expression called the *determinant* of a 3×3 pattern of coefficients: it is enough to say at this stage that for the echelon pattern the determinant is equal to the product $a_1b_2c_3$.

EXERCISE 5b

(1) Using the above method, solve the slightly modified set of equations indicated:

$$\begin{array}{rrr|r} 1 & 2 & -1 & -1 \\ 2 & 7 & 5 & -8 \\ 7 & 6 & 2 & 9 \\ \hline \end{array}$$

Give also the determinant of your final set of equations.

(2) Reduce the set of equations given below to an echelon form. Show that the determinant of this form is zero, but that in this case *any* value may be assigned to z; i.e. the failure is only a failure to get a unique solution for x, y and z.

$$\begin{array}{rrr|r} 1 & 2 & 7 & 12 \\ 2 & 7 & 6 & 11 \\ 1 & -1 & 15 & 25 \end{array}$$

Take $z = 2$ and determine the corresponding values of x and y. Repeat with another value of z, of your own choice.

(3) Using two facts of 3-dimensional co-ordinate geometry, (*i*) that a condition such as $x+8y-5z = -1$ determines a plane locus and (*ii*) that *in general* two planes meet in a straight line, give a geometrical interpretation to the relationships of the planes in questions (1) and (2) above.

(4)-(9) In the following questions reduce the set of coefficients to echelon form, giving (*i*) at each stage a complete set of 3 rows of coefficients, and (*ii*) a clear indication of how this set is derived from the previous set. A change in the order of the rows may be advisable at the start or at any subsequent stage. (The solution of the equations is, as it were, a by-product of this work.)

(4) $$\begin{array}{rrr|r} 2 & -3 & 5 & 10 \\ 4 & 7 & -2 & -5 \\ 2 & -4 & 25 & 31 \end{array}$$

(5) $$\begin{array}{rrr|r} 2 & -3 & 5 & 13 \\ -4 & 6 & 18 & 2 \\ 6 & 3 & 0 & 0 \end{array}$$

(6) $$\begin{array}{rrr|r} 2 & 0 & 9 & 7 \\ 6 & 1 & 3 & -8 \\ 0 & -5 & 8 & 33 \end{array}$$

(7) $$\begin{array}{rrr|r} 5 & 7 & 3 & 0 \\ 1 & 2 & 9 & 9 \\ 8 & 10 & 7 & 1 \end{array}$$

(8) $$\begin{array}{rrr|r} 3 & -4 & 1 & 2 \\ -5 & 6 & 10 & 7 \\ 7 & -10 & 5 & 6 \end{array}$$

(9) $$\begin{array}{rrr|r} 3 & 2 & -5 & 1 \\ 0 & 4 & 3 & 1 \\ 6 & 4 & -10 & 2 \end{array}$$

chapter 6

THE FOUNDATIONS SURVEYED

We now proceed to put into formal algebraic shape the work which has been done piece by piece up to this point. (The student need not at first reading attempt to memorise points of general algebraic structure, but will need to appreciate clearly the concept of vector space and its associated terms.)

6.1 The reader is reminded that it is of the nature of a formal algebra to have:

(*i*) a set of elements, which may be finite or not;

(*ii*) a law of composition by which any pair of elements α, β of the set will give uniquely a third element γ, also of this set: we may express this as $\alpha * \beta = \gamma$ (remembering that α, β do not have to be distinct, so that $\alpha * \alpha$ is also a member).

Some of the algebraic systems which we considered had one or more of the following additional properties:

(*iii*) To contain an element e such that for *every* element α of the set $\alpha * e = e * \alpha = \alpha$.

[We call such an element the *identity* element of the system.]

(*iv*) To have for every element α a mate α' such that when α is combined with its mate we arrive at e; i.e. $\alpha * \alpha' = \alpha' * \alpha = e$, in which case α' is the *inverse* of α and written α^{-1}. (α^{-1} does not have to be distinct from α: sometimes an element is its own inverse.)

(*v*) To have a rule governing the combination of any three elements; viz. for p, q, r in that order it does not matter whether we work out $(p * q) * r$ or $p * (q * r)$, the resulting element is the same. This rule or property is called *associativity*: it looks to be the most subtle of the properties but it is found to be satisfied by a wide range of structures.

We can regard (*i*) and (*ii*) as the essential minimum which qualifies a system to be 'an' algebra: the possession of all the rest gives the algebra a particularly tidy and useful form. Such a system is called a GROUP.

Vectors of a particular type and size (number of components) form a group; we shall note how they fit into the above pattern, taking 2-vectors as our example. (They *can* be defined geometrically, but it is better to use algebraic definitions, in this case column 2-vectors, and illustrate with geometrical *pictures*.)

In detail for this system:

(*i*) The typical element is $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ where a and b are real numbers.

The picture is of a displacement which increases the x-co-ordinate by a and the y-co-ordinate by b, on a plane with a pair of given axes. The set is infinite.

(*ii*) The law is 'addition' as defined earlier: it is commutative.

(*iii*) $\mathbf{e} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

(*iv*) Element $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ has inverse $\begin{pmatrix} -a \\ -b \end{pmatrix}$.

(*v*) Associativity is obviously obeyed, since it is true for addition of the numbers which form the components.

There are clear geometrical pictures for each of the above properties, as the reader should try out for himself.

EXERCISE 6a

(1) If a and b belong to the set of integers, let $a * b = a+b+1$.

(*a*) Which integer e has the property that $a * e = e * a = a$ for every element a?

(*b*) Is there an inverse of a for this law?

(*c*) Do we have associativity?

(*d*) Is the system a group for this law of combination?

Answer the same questions for $a \circ b = a+b+ab$.

(2) Consider the numbers 0, 1, 2, 3, 4, 5 under addition modulo 6. (See exercise 1c, question 1, e.g. $5 \oplus 4 = 3$.) Do they form a group? If so, give the inverse of each element; if not, say why not.

(3) Consider the same numbers as in question 2 under *multiplication* modulo 6.

The reader should return to the exercises of chapter 1 and work any questions that were omitted then.

6.2 We have looked at the group properties (*i*) to (*vi*) in our vector system, but like some other interesting systems it has further properties beyond this. (The familiar real-number system has of course *two* laws of combination: multiplication gives a further structure which all but satisfies rules (*i*) to (*vi*) a second time, zero being the exception because it has no inverse for multiplication.)

A vector system also incorporates numbers as multipliers in the way described earlier, so that from any vector $\mathbf{u}$ we may form another, say $p\mathbf{u}$, and from $\mathbf{v}$ we can form $q\mathbf{v}$. Furthermore, just as we could add $\mathbf{u}$ and $\mathbf{v}$, we can add $p\mathbf{u}$ and $q\mathbf{v}$; i.e. from $\mathbf{u}$ and $\mathbf{v}$ as building-bricks we can form an infinite number of new elements of the form $p\mathbf{u}+q\mathbf{v}$ and these are all members of the set. This is, so to speak, a manufacturing process for vectors of the system. But we have next to consider whether *every* vector of the system can be manufactured in this way; i.e. given a vector $\mathbf{w}$ of the system, do numbers p and q exist so that $\mathbf{w} = p\mathbf{u}+q\mathbf{v}$? We have already considered this as a practical problem in connection with the solution of simultaneous linear equations, and derived the condition, viz. that the vectors $\mathbf{u}$, $\mathbf{v}$ should be linearly independent.

This is the point at which to introduce three new technical terms. First, a vector system as described above is called a VECTOR SPACE†—a term which shows the origin of the idea, but which might cause the reader to forget that such systems exist for vectors of order greater than three.

Secondly, if any two 2-vectors $\mathbf{u}$ and $\mathbf{v}$ are taken as bricks, then the set of all vectors which are built out of $\mathbf{u}$ and $\mathbf{v}$ is said to be the space SPANNED by $\mathbf{u}$ and $\mathbf{v}$.

Thirdly, if these 2-vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent they are said to form a BASIS for this vector 2-space, and we have seen in chapter 5 that three linearly independent 3-vectors provide a basis for a 3-space. The geometry of 3-dimensions, developed by using Cartesian axes Ox, Oy, Oz, can be re-stated in terms of vectors, with, as basis, the unit vectors $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, a unit vector parallel to the x-axis, together with $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Then the position of a point P is precisely defined by the vector $\overline{OP}$ where $\overline{OP} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$.

† The full formal definition of a vector space is held over until the end of the chapter.

A certain amount of care must be taken over these words, and some examples will serve to illustrate. It must not be supposed that a vector-space comprising 3-vectors necessarily needs three vectors as a basis. There is a vector space for which **i** and **j** alone are a basis; and even one for which **i** alone is a basis. For this last vector space, all its vectors are parallel to the x-axis of co-ordinates.

Another example of a vector space is the space which has $\begin{pmatrix}2\\1\end{pmatrix}$ or $\begin{pmatrix}3\\1{\cdot}5\end{pmatrix}$ or $\begin{pmatrix}4\\2\end{pmatrix}$ or $\begin{pmatrix}2/3\\1/3\end{pmatrix}$ as a basis; in fact the vector space spanned by the four vectors given is the same as that for which $\begin{pmatrix}2\\1\end{pmatrix}$ is a basis.

It will be apparent that there is no precision in the idea of a basis, unless we choose a *minimum* number of vectors for it. If the vectors comprising a basis were linearly dependent, then one could be dispensed with.† It is essential that the vectors of a basis should be *linearly independent*.

The number of vectors required for a basis is called the DIMENSION of the space. In all examples where the dimension is less than or equal to 3 this agrees with our intuitive geometric ideas.

There is no need for the basis-vectors to be of the type **i**, **j** and **k** shown above, though to have three unit vectors which are mutually perpendicular is an enormous advantage. We have already seen that an echelon set will do, e.g. $\begin{pmatrix}a_1\\0\\0\end{pmatrix}\begin{pmatrix}p\\b_2\\0\end{pmatrix}\begin{pmatrix}q\\r\\c_3\end{pmatrix}$. These may be used to build up any vector of the form $\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}$ provided, of course, $a_1b_2c_3 \neq 0$, and hence act as a basis.

EXERCISE 6b

(1) Find numbers λ, μ and ν such that $\lambda\begin{pmatrix}3\\2\end{pmatrix} + \mu\begin{pmatrix}5\\3\end{pmatrix} + \nu\begin{pmatrix}4\\3\end{pmatrix} = 0$.

Show that any 2-vector may be written as a combination of these three vectors, i.e. that they *span* the space comprising 2-vectors. Show further that any one of the three may be omitted to form a linearly independent pair of vectors which will act as a *basis* for the space of 2-vectors.

(2) Show that the vectors $\begin{pmatrix}3\\1\end{pmatrix}$, $\begin{pmatrix}6\\2\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ span the space of 2-vectors, but that it is not possible to find a basis for 2-space by omitting *any* one of them.

(3) The following vectors span a vector space:

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}\begin{pmatrix}2\\0\\1\\0\end{pmatrix}\begin{pmatrix}1\\0\\3\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}.$$

How many are needed to form a basis for this space, and what will be its dimension? Does it matter which vectors are omitted to form a basis?

† We are not stating that *any* one can be dispensed with.

6.3 The algebraic structure called a vector space has a very wide range of applications. A topical example is in colour television, where the three-fold colour information for every spot on the screen (initially observed by the camera as intensities of red, green and blue) is expressed and transmitted in terms of a new basis (white, and two 'chromaticities').

We shall take for discussion a structure within mathematics, viz. the set of all quadratic expressions in x with integral coefficients. For example, take $2x^2-5x+3$, $-7x^2$, $5x-4$ $(=0x^2+5x-4)$. The general form is of course ax^2+bx+c, where a, b, c can have all integral values not excluding zero.

It is clear:

(*i*) that such an expression is uniquely determined by its triad of coefficients e.g. (a_1, b_1, c_1) or (a_2, b_2, c_2);

(*ii*) that a law of combination exists, viz. addition, where $(a_1, b_1, c_1)+(a_2, b_2, c_2) = (a_1+a_2, b_1+b_2, c_1+c_2)$ and the result is a member of the set—i.e. the set is closed under this law;

(*iii*) that in view of our definition there is a 'zero' quadratic $(0, 0, 0)$;

(*iv*) that a quadratic (a, b, c) has an inverse $(-a, -b, -c)$;

(*v*) that addition is associative for these elements, since it is for the component integers.

This shows that we have a GROUP but not yet a vector space.

If Q is a quadratic of this system, then so also is nQ where n is an integer. More generally, $lQ_1+mQ_2+nQ_3$ is also a member of the system. We could in fact take $Q_1 = (1, 0, 0)$, which is the quadratic

x^2+0x+0; $Q_2 = (0, 1, 0)$ which is $0x^2+x+0$; and $Q_3 = (0, 0, 1)$ which is $0x^2+0x+1$. This would provide a basis on which we could build up the whole system. If, however, we think of this in geometrical terms, with these as position-vectors for a set of points, we find ourselves left with the points with integral coefficients. Not only does this offend our idea of the term space, but it is in fact contrary to the technical definition of a real vector space (see section 6.4) which requires any real multiple of a vector to be also a vector of the space.

If we start again and define our polynomial as having real coefficients (a, b, c) we have a vector space which obeys all the rules. Since integers are real numbers we may still take Q_1, Q_2, Q_3 as defined above to be our basis: for most purposes it is the simplest basis.

We can handle in concrete fashion vector spaces of higher dimension than three by considering polynomials of higher degree; e.g. we may go up to 4th degree in the form (a, b, c, d, e). Notice that the quadratic expression x^2 belongs to this space as $q_1 = (0, 0, 1, 0, 0)$, and we also have $q_2 = (0, 0, 0, 1, 0)$ and $q_3 = (0, 0, 0, 0, 1)$ as members. With these as basis we span a subspace $lq_1+mq_2+nq_3$. (A subspace is defined as a subset of the full space which itself satisfies all the conditions of a space.)

We conclude this chapter with two algebraic sections, one to give in full rigour the definition of a real vector space and the other to develop more algebraic ideas (relations and mappings) which will be required for later work.

6.4 *Definition of a real vector space V*

V is a set of elements which form a *group* under addition (with zero written $\mathbf{0}$). Further properties required are as follows:

If $\mathbf{x} \in V$ and r, s are real numbers:

(*i*) $r\mathbf{x} \in V$
(*ii*) $(rs)\mathbf{x} = r(s\mathbf{x})$
(*iii*) $(r+s)\mathbf{x} = r\mathbf{x}+s\mathbf{x}$
(*iv*) $r(\mathbf{x}+\mathbf{y}) = r\mathbf{x}+r\mathbf{y}$
(*v*) $1.\mathbf{x} = \mathbf{x}$
(*vi*) $0.\mathbf{x} = \mathbf{0}$

The reader may, with a picture (or a column-vector) in his mind, think that part of the above definition is unnecessary, e.g. (*ii*). If the algebraist can construct other systems which have (*i*), (*iii*), (*iv*), (*v*), (*vi*), but not (*ii*), then (*ii*) is an *essential* part of the definition. It is as if we define a teapot as having a spout, a handle and a lid: all are essential to the definition because a kind of pot could be made without having all three and we should not accept such a pot as being a teapot.

The student would, however, be justified in dropping (*vi*), because it can be proved from (*iii*) and the group character of V. (Write $r = 1$ and $s = -1$.) We include it for clarity.

EXERCISE 6c

Show that the elements $2x+1$, $x-\frac{1}{2}$, $x+\pi$, $2x-1$ span a vector space which consists of all linear expressions in one variable x (i.e. of the form $px+q$ where p and q are real numbers). Show also that (*i*) one pair and (*ii*) any three of the given four elements are linearly dependent. Express $x-1$ in terms of $2x+1$, $x-\frac{1}{2}$ as a basis. Suggest (*iii*) a simpler basis for this vector space, (*iv*) a basis for the vector space consisting of all quadratic expressions px^2+qx+r. Are x^2, $x-2$, $3x^2-2x+4$ linearly independent elements in this space?

Relations

6.5 'Relation' is one of the many everyday English words which are used in mathematics as (precisely defined) technical terms. A relation is defined for two sets X, Y (which are best visualised as different though they need not be so).

Definition: A relation ρ between sets X, Y is a defined set of ordered pairs (x, y). A simple example is the set of two-course meals {entrée, sweet} which are chosen by a party of diners from a given menu, i.e. from a specified set of entrées and a specified set of sweets. It will be seen that if there are m entrées and n sweets, a relation defined upon them could have any number of members from zero (i.e. a null relation) to the set of all possible pairs, which has mn members. The latter is called the *product set* and written $X \times Y$.

It should be noted that throughout this work the stated order is important; e.g. if any diner were to have his sweet first his meal would not belong to the relation X, Y but to a different relation, Y, X. The full product-set $Y \times X$ comprises all the possible pairs y, x, i.e. {sweet, entrée}.

Relations already familiar to the reader, though perhaps not by that name, involve ordered pairs of real numbers. Each member of such a

relation is a number-pair, e.g. $(\pi, -1)$: it may be visualised as a point on graph-paper. This particular number-pair belongs to the relation $\{(x, y) \mid y = \cos x\}$, the vertical stroke being read as 'such that'. Every kind of graph or 'scattergram' exhibits such a relation; and all are subsets of the vast structure consisting of all possible pairs (x, y). This set is often written $R \times R$ since it pairs every member with every member of the set R of real numbers.

In general, relations in the set $R \times R$ are defined by some rule of calculation which determines those values of y which are paired with a given value of x and vice versa; $y = \cos x$ is an example of such a rule.

Relations are classified in two different ways: (*a*) according to whether certain of the available x's (or y's) do not enter into it, e.g. for $y = \cos x$ all values of x enter into the relation but only y-values in the range $-1 \leqslant y \leqslant 1$; (*b*) according to the kind of tie-up between a typical x-value and the y-values with which it is paired: this example is called a *many-one* relation, because many values of x are paired with the same value of y. For example $(-5\pi, -1)$ $(-3\pi, -1)$ $(-\pi, -1)$ $(\pi, -1)$ $(3\pi, -1)$.

Other types of relation are one-one, one-many, many-one, and many-many; the product set is an extreme example of a many-many relation. The word 'many' is included in the title provided some of the elements conform to the pattern described, e.g. for real numbers $\{(x, y) \mid x^2 = y\}$ qualifies as many-one although for $x = 0, y = 0$ it is one-one. In fact one isolated case in which two x-values are linked with a single y would be enough for a relation to qualify as many-one.

A many-one relation into which every x enters is called a *function* or a *mapping* of the set X into the set Y. This is a type of relation with which we shall have much to do. An even stricter type of relation is a *one-to-one correspondence*: in this case every x and every y enters exactly once, each having a mate in the other set. We have seen a variety of examples of this situation, which in a geometrical context is called a *reversible transformation*.

EXERCISE 6d

(1) For each of the following relations on sets X, Y of real numbers, state:

(*i*) whether any values of x or of y do not enter into it;
(*ii*) whether it is one-one, many-one, one-many or many-many.

(*a*) $\{(x, y) \mid x = y^2\}$ (*b*) $\{(x, y) \mid x^2+y^2 = 1\}$
(*c*) $\{(x, y) \mid xy = 0\}$ (*d*) $\{(x, y) \mid y > x+1\}$
(*e*) $\{(x, y) \mid x+y = 2\}$

(2) For each of the following relations on sets P, Q of integers state:

(*i*) whether any values of p or of q do not enter into it;
(*ii*) whether it is one-one, many-one, one-many or many-many.

(*a*) $\{(p, q) \mid p^2+q^2 = 25\}$ (*b*) $\{(p, q) \mid p^2-q^2 = 1\}$
(*c*) $\{(p, q) \mid p = 2q\}$ (*d*) $\{(p, q) \mid q = \sin 2\pi p\}$
(*e*) $\{(p, q) \mid p^2+q^2 = 2pq+1\}$ (*f*) $\{(p, q) \mid p^2+q^2 = 0\}$

(3) Discuss as above, for sets $\{\mathbf{a}, \boldsymbol{\alpha}\}$ of 2-vectors with real components, the relations defined by:

(*a*) $\boldsymbol{\alpha} = \mathbf{a}+\mathbf{k}$, where $\mathbf{k} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

(*b*) $\boldsymbol{\alpha} = 7\mathbf{a}$

(*c*) $\boldsymbol{\alpha} = a_2\mathbf{i}+a_1\mathbf{j}$, where $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$

(*d*) $\boldsymbol{\alpha} = \mathbf{a}$ where $a_1{}^2+a_2{}^2 = 1$, and $\boldsymbol{\alpha} = \mathbf{0}$ otherwise.

chapter 7

SOME VECTOR GEOMETRY: LENGTH, DIRECTION COSINES, SCALAR PRODUCT

7.1 Let us suppose that we are considering 2-vectors, in a plane,† and we have decided upon a basis for this space. Clearly, we want to know as much about the base vectors as we can. Two factors spring to mind from the geometrical picture: first, the length of the vectors; and secondly, the directions of (and the angle between) the vectors.

What about the length? If we consider the vector $\begin{pmatrix}5\\12\end{pmatrix}$ we may rewrite $\begin{pmatrix}5\\12\end{pmatrix} = 5\mathbf{i}+12\mathbf{j}$ where $\mathbf{i}$ and $\mathbf{j}$ are vectors of unit length parallel to the x and y axes.

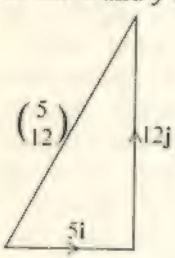


FIG. 22

This automatically constructs a right-angled triangle for us, which gives the length of $\begin{pmatrix}5\\12\end{pmatrix}$ as 13. Similarly the length of $\begin{pmatrix}a_1\\a_2\end{pmatrix}$ is $\sqrt{a_1^2+a_2^2}$, and a 3-dimensional figure shows us that the length of $\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}$ can be found.

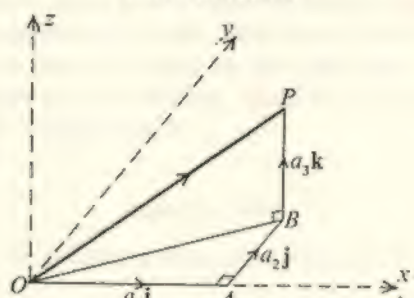


FIG. 23

† We write this because we could have a set of 3-vectors in a plane: the order of the vectors and the dimension of the space are not necessarily the same.

We have $OB^2 = OA^2+AB^2 = a_1^2+a_2^2$.

Also, $OP^2 = OB^2+PB^2 = a_1^2+a_2^2+a_3^2$.

$$\text{Hence, length } \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = \sqrt{a_1^2+a_2^2+a_3^2}.$$

We are in danger of allowing the geometry to run the algebra, so we now *define* the MODULUS of a vector $\mathbf{a} = \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}$ (effectively the length) as $\sqrt{a_1^2+a_2^2+a_3^2}$, and to denote this we write the symbol $\mathbf{a}$ between vertical lines, thus: $|\,\mathbf{a}\,|$.

Hence, $|\,\mathbf{a}\,| = \sqrt{a_1^2+a_2^2+a_3^2}$.

EXERCISE 7a

(1) Find the modulus of each of the following vectors and sketch where possible.

(*a*) $\begin{pmatrix}5\\0\end{pmatrix}$ (*b*) $\begin{pmatrix}-5\\0\\0\\0\end{pmatrix}$ (*c*) $\begin{pmatrix}3\\4\end{pmatrix}$ (*d*) $\begin{pmatrix}0\\-3\\4\end{pmatrix}$

(*e*) $\begin{pmatrix}12\\3\\-4\end{pmatrix}$ (*f*) $\begin{pmatrix}-12\\0\\3\\4\end{pmatrix}$ (*g*) $\begin{pmatrix}\cos\theta\cos\phi\\\cos\theta\sin\phi\\\sin\theta\end{pmatrix}$

(*h*) $\begin{pmatrix}\sqrt{3}\\\sqrt{3}\\-\sqrt{3}\end{pmatrix}$ (*i*) $\begin{pmatrix}-1\\-1\\-1\end{pmatrix}$ (*j*) $\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{-1}{\sqrt{2}}\\0\end{pmatrix}$

(2) An octahedron has vertices at the points $(0, 0, \pm\sqrt{2}a)$ $(\pm a, a, 0)$ $(\pm a, -a, 0)$. Show that all its edges are equal in length, and express their directed lengths as 3-vectors referred to the axes of co-ordinates. (The *sense* given to each line should be shown on a figure.)

(3) A gun fires a projectile with velocity 4,000 ft/sec. at an elevation of 60°, on a horizontal bearing 030° (i.e. N 30° E). Express the velocity as a vector in terms of **e**, **n**, **u**, unit vectors east, north and upwards. If the recoil of the gun imposes an extra velocity of 40 ft/sec. horizontally in direction 210°, state the actual velocity of the shell as a vector.

(4) The electric field at a point inside a cathode-ray tube is given by its components in the x-, y-, and z-directions, measured in volts per cm. At a certain instant it changes from $\mathbf{F}_1 = (5, -3, 11{\cdot}5)$ to $\mathbf{F}_2 = (-3, 1, 12{\cdot}5)$. Show that the magnitude (modulus) of the field is unchanged. Calculate the modulus of $(\mathbf{F}_1 - \mathbf{F}_2)$ and hence or otherwise show that the field direction has turned through an angle of nearly 41°.

(5) ABC is an equilateral triangle which forms the base of a tetrahedron $OABC$ in which the lengths of OA, OB, OC are not necessarily equal. Show that the centroids of the faces OAB, OBC, OCA form an equilateral triangle. Find the ratio of the area of this triangle to that of ABC.

(6) ABC is a triangle with centroid at O, and X is any point in its plane. Prove that $XA^2+XB^2+XC^2 = 3XO^2+AO^2+BO^2+CO^2$.
(*Hint*: Take **a**, **b**, **c** as position-vectors with respect to O, and use $\mathbf{a}+\mathbf{b}+\mathbf{c} = \mathbf{0}$.)

(7) Prove that the quadrilateral formed by joining the midpoints of the sides of a given plane quadrilateral is a parallelogram.

Investigate whether the above result is true if the original quadrilateral is skew.

7.2 *Direction of a vector*

Specifying direction begins to be tricky only in 3-dimensions, and we shall therefore start there.

We have no hesitation in saying that the vectors $\begin{pmatrix} 3 \\ 12 \\ 4 \end{pmatrix} \begin{pmatrix} 6 \\ 24 \\ 8 \end{pmatrix}$ have the same direction; and in a more general case that $\mathbf{a}$, $k\mathbf{a}$ have the same direction (provided k is positive).† We could therefore specify a direction in either of two ways:

† We have in each case a most elementary form of linear dependence.

(*i*) We could just give the components in order, for *any* vector in that direction: in the above case we should prefer to do it in lowest terms as 3, 12, 4. These are then called *direction numbers* for this direction.

(*ii*) We could scale them down so that we get a *unit* vector in the required direction. In this example our given vector is of modulus (i.e. length) $\sqrt{3^2+12^2+4^2} = 13$: we would therefore divide all the numbers by 13 and specify the direction by $\frac{3}{13}, \frac{12}{13}, \frac{4}{13}$. (The reader should verify that these numbers, squared and added, give unity.)

Such numbers are called *direction-cosines*, because they are in turn $\cos\theta_1$, $\cos\theta_2$, $\cos\theta_3$ where $\theta_1, \theta_2, \theta_3$ are the angles which the directed line makes with the co-ordinate axes Ox_1, Ox_2, Ox_3. The reader should convince himself, if necessary by construction of a model, that for the unit-vector $\begin{pmatrix} 3/13 \\ 12/13 \\ 4/13 \end{pmatrix}$ these trigonometrical results are true.

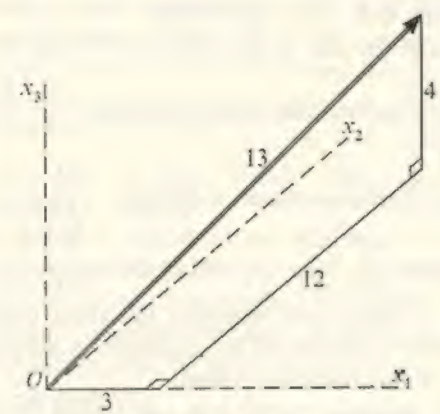


FIG. 24

The diagram shows the situation in perspective.

Worked example

O is the centre of a cube with vertices $ABCDA'B'C'D'$. Axes are taken through O perpendicular to the faces as shown in figure 25.

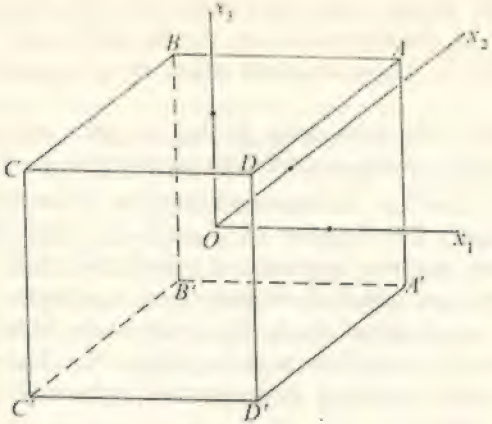


FIG. 25

Required to find: the direction-cosines of the eight directions $\overline{OA}$, $\overline{OB}$, $\overline{OC}$, $\overline{OD}$, $\overline{OA'}$, $\overline{OB'}$, $\overline{OC'}$, $\overline{OD'}$.

To solve, we take the side of the cube as 2. Then the vector $\overline{OA}$ is $\begin{pmatrix}1\\1\\1\end{pmatrix}$ of length $\sqrt{3}$ and the direction-cosines are $\frac{1}{\sqrt{3}}$ $\frac{1}{\sqrt{3}}$ $\frac{1}{\sqrt{3}}$. Since B is $(-1, 1, 1)$, the direction-cosines of $\overline{OB}$ are $\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$.

Similarly, C gives $\frac{-1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$ and D gives $\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$.

In every one of these four instances the third cosine is positive, corresponding to the fact that all these directions make an acute angle with Oz (or Ox_3 if this notation is used).

With reference, however, to the first axis (called Ox or Ox_1 as preferred), they differ; e.g. OB makes an obtuse angle, and this is shown by its first direction-cosine $\left(\frac{-1}{\sqrt{3}}\right)$.

The direction-cosines of OA' are those of OA with signs reversed, and similarly for OB', OC' and OD'.

EXERCISE 7b

(1) Find the direction-cosines of the lines joining:

(*a*) $(2, 4, -3)$ to $(1, 2, -1)$
(*b*) $(3, -1, 2)$ to $(5, 2, -4)$
(*c*) $(-1, 3, -6)$ to $(2, -1, 6)$.

(2) $ABCD$ is a face of a cube, O is the centre, and $A'B'C'D'$ is the opposite face, AOA', BOB' etc. being straight lines. Give the direction-cosines of $\overline{OA}$, $\overline{OB}$, $\overline{OC}$, $\overline{OD}$ for each of the following choices of axes:

(*a*) P, Q, K are the midpoints of the faces $ABCD$, $BCA'D'$, $ABD'C'$ respectively; and $\overline{OP}$, $\overline{OQ}$, $\overline{OK}$ are taken as Ox_1, Ox_2, Ox_3.
(*b*) S, T are the midpoints of AD, BC respectively; and OS, OT, OK are taken as Ox_1, Ox_2, Ox_3.

(3) A wooden block is constructed in the shape of a pyramid on a square base $ABCD$, the slant edges VA, VB, VC, VD being equal in length to the sides of the base. The block is turned upside down so that AB is horizontal and eastwards, and BC is horizontal and northwards. Give the direction cosines of all the edges, referred to axes eastward (Ox_1), northward (Ox_2) and upward (Ox_3).

(4) With respect to rectangular axes through O, A is $(0, -8, 6)$, B is $(6, 0, -8)$, C is $(3, -4, -1)$, D is $(6, -8, -2)$. Find the direction-cosines of vectors $\overline{OA}$, $\overline{OB}$, $\overline{CA}$, $\overline{CB}$, $\overline{DA}$, $\overline{DB}$; and comment on the results with the aid of a diagram.

(5) An exposed object is subject to a wind force which is expressible in lbf (pound-force) with respect to axes east, north and upwards as a row-vector (24, 32, 9). Find the magnitude of the force, and the cosines of the angles which it makes with the axes. Express also as a row-vector a force of 200 lbf in the same direction as before.

(6) (Data as in question 5.)
What is the inclination of the wind-force to the horizontal *plane*? If this force is equivalent to a single horizontal force **H** lbf together with a vertical force **V** lbf, give (*a*) the magnitudes H, V and (*b*) the *bearing* of the force **H**, measured in standard fashion (clockwise from the north direction).

7.3 *Angle between two vectors in 3-dimensions*

If the directions of two vectors are specified, whether by their direction-cosines or direction numbers, the angle between them is definite and it

must be possible to calculate it. A simple plan will be to consider two vectors $\overline{OA} = \mathbf{a}$ and $\overline{OB} = \mathbf{b}$, not necessarily of unit magnitude, where

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

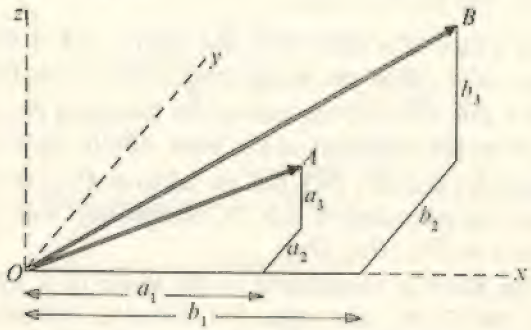


FIG. 26

We know all the sides of triangle AOB, and hence we can use the cosine formula to find the angle ϕ ($= A\hat{O}B$) between the directions of the vectors $\mathbf{a}$, $\mathbf{b}$:

$$AB^2 = OA^2+OB^2-2OA.OB\cos\phi.$$

But $AB^2 = (a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2$ by Pythagoras's Theorem $= (a_1^2+a_2^2+a_3^2)+(b_1^2+b_2^2+b_3^2)-2(a_1b_1+a_2b_2+a_3b_3)$;

and the other expression

$$= (a_1^2+a_2^2+a_3^2)+(b_1^2+b_2^2+b_3^2)-2\,|\,\mathbf{a}\,||\,\mathbf{b}\,|\cos\phi.$$

Hence $|\,\mathbf{a}\,|\,.\,|\,\mathbf{b}\,|\cos\phi = a_1b_1+a_2b_2+a_3b_3$; a result which is often written using a, b for the moduli, i.e. lengths OA, OB:

$$ab\cos\phi = a_1b_1+a_2b_2+a_3b_3.$$

From this equation we can always derive the angle between two vectors, and in particular between two *directions*, if we are given them as unit-vectors $\begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$, $\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}$. In this case $\cos\phi = l_1m_1+l_2m_2+l_3m_3$, the cosine of the required angle being given in terms of the direction-cosines of the two directions. The reader should verify this result also by pure geometry: see exercise 7c, question 13.

Algebraic formulation

In accordance with our rule of not letting the geometry run the algebra, but only illustrate it, we proceed by defining in algebra:

(*i*) The *inner product* (*or scalar product*) of two 3-vectors $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ is defined as $a_1b_1+a_2b_2+a_3b_3$, with a similar definition for vectors of any order (so long as $\mathbf{a}$ and $\mathbf{b}$ are of the *same* order, e.g. for two-vectors $a_1b_1+a_2b_2$).

(*ii*) Two vectors are said to be *orthogonal* if their inner product is zero: for 3-vectors this would give us $\cos\phi = 0$, so that the vectors are at right angles.†

(*Note*: The formation of an inner product from two given vectors $\mathbf{a}$, $\mathbf{b}$ of our system is an example of a law of combination; but notice that the product is a number, not a vector. This is therefore a process not in the pure algebra of vectors, but only in an extended system which has numbers as well as vectors.)

Since the term 'product' has been used, the reader will not be surprised to see the notation $\mathbf{a}.\mathbf{b}$ for $a_1b_1+a_2b_2+a_3b_3$.

Thus $|\,\mathbf{a}\,||\,\mathbf{b}\,|\cos\phi = \mathbf{a}.\mathbf{b}$, and in particular $\cos\phi = \hat{\mathbf{l}}.\hat{\mathbf{m}}$, where $\hat{\mathbf{l}}$, $\hat{\mathbf{m}}$ are unit-vectors (as the 'hat' notation indicates). The components of $\hat{\mathbf{l}}$ are, of course, the actual direction-cosines l_1, l_2, l_3.

The condition for *orthogonality* of vectors, whatever their moduli, is $\mathbf{a}.\mathbf{b} = 0$; but notice that in geometry $\mathbf{a}.\mathbf{b} = 0$ only implies that the vectors are at right angles if neither has zero modulus.

(*Note*: Owing to the notation the scalar product is sometimes known as the 'dot product' of vectors. It is clearly commutative. The dot must never be omitted.)

Worked examples

(*a*) A tetrahedron has as horizontal base an equilateral triangle BCD of side $3\sqrt{3}$, and the edges AB, AC, AD are all 5. It is placed on a table with CD lying northward and B towards the east. O is the centre of the base. The axes Ox_1, Ox_2, Ox_3 being east, north and upwards,

† The term orthogonal, as an algebraic term, can be used for vectors of order greater than 3.

give direction numbers for $\overline{GA}$, $\overline{GB}$, $\overline{BA}$, $\overline{CD}$, where G is a point one unit vertically above the centre of the base. Hence (*i*) verify that $\overline{BA}$, $\overline{CD}$ are at right angles, (*ii*) find $A\widehat{G}B$. Show that the sum of vectors from G to the vertices is zero.

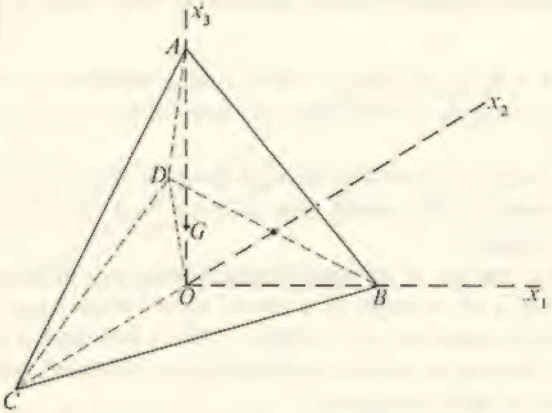


FIG. 27

Solution: Take O as centre of the base. Then $OB = OC = OD = 3$ and $AO = 4$.
Position-vectors with O as origin are:

$$\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\overline{GA} = -\mathbf{g}+\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}, \qquad \overline{GB} = -\mathbf{g}+\mathbf{b} = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$$

$$\overline{BA} = -\mathbf{b}+\mathbf{a} = \begin{pmatrix} -3 \\ 0 \\ 4 \end{pmatrix},$$ $\overline{CD}$ is parallel to Ox_2 and its direction numbers are 0, 1, 0.

We see that
$\overline{BA}.\overline{CD} = (-3)(0)+(0)(1)+(4)(0) = 0 \Rightarrow$ orthogonality,
and $\overline{GA}.\overline{GB} \cos\phi = (0)(3)+(0)(0)+(3)(-1) = -3$;
i.e. $3\sqrt{10}\cos\phi = -3$.

Therefore $\cos\phi = \dfrac{-1}{\sqrt{10}}$.

(*b*) Find direction numbers for a direction which is orthogonal both to

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \text{ and to } \mathbf{b} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}.$$

Solution: We require numbers p, q, r such that $p+2q-r = 0$ and $3p-4q+r = 0$.
By addition, $4p-2q = 0 \Rightarrow q = 2p$.
Hence $r = p+2q = 5q$.
We may take $p = 1$ and get numbers 1, 2, 5.

If we write $\mathbf{c} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$ it is clear that $\left.\begin{matrix} \mathbf{a}.\mathbf{c} = 0 \\ \text{and } \mathbf{a}.\mathbf{b} = 0 \end{matrix}\right\}$ check.

EXERCISE 7c

(1) A missile is said to 'vector' on a target when it has full information about the position of the target with respect to itself as origin. At a certain instant this information is, with respect to certain axes, the vector $\begin{pmatrix} 8 \\ 24 \\ 6 \end{pmatrix}$ and the missile's velocity is given by the vector $\begin{pmatrix} 15 \\ 20 \\ 0 \end{pmatrix}$.

Calculate the angle between these two directions.

(2) A vertical mast is supported by three equal stay-wires (which may be treated as straight), all attached to a point P 100 ft up on the mast, and on the ground to points A, B, C all 50 ft from the foot F of the mast and symmetrically around it. For safety the wire PA is now replaced by another longer one PD, such that PD is perpendicular to both the other wires. How long is FD? (*Hint*: With P as origin, take axis Ox_3 downwards and Ox_1, Ox_2 horizontally, the latter in the plane of symmetry.)

The reader should check his solution by the geometry of the figure projected onto the meridian plane.

(3) Given $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{j}, \mathbf{k}$ similarly:

(*i*) Prove $\mathbf{a}.\mathbf{a} = |\mathbf{a}|^2$; $\mathbf{i}.\mathbf{j} = 0$; $\mathbf{i}.\mathbf{i} = 1$.

(*ii*) Prove $\mathbf{a}.\mathbf{i} = a_1$; and state similar results involving $\mathbf{k}$.

(4) Given $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$, $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}$, verify that $\mathbf{a}.(\mathbf{b}+\mathbf{c}) = \mathbf{a}.\mathbf{b}+\mathbf{a}.\mathbf{c}$. Establish this result (distributivity for dot product over vector addition) in general for 3-vectors.

(5) Two boys compared their marks after examinations. They were as follows:

	English	*French*	*Mathematics*	*Science*
Boy A	80%	50%	70%	60%
Boy B	60%	40%	80%	80%

They wished to decide (*i*) which had done better, and (*ii*) to what extent their abilities lay in different directions. Treating rows of marks as vectors (**a**, **b**) of a space, try to obtain answers to these questions. Discuss the validity of the assumptions made in such a treatment.

(6) In mechanics the work done by a constant force **F** which moves its point of application by a displacement **s** is defined either as $\mathbf{F}.\mathbf{s}$ or as $|\mathbf{F}||\mathbf{s}|\cos\phi$ where ϕ is the angle between the directions of **F** and **s**. Why can these definitions be regarded as equivalent? Show that the work done is the (scalar) sum of the works done by the component forces, i.e. by $F_1\mathbf{i}$ moving through $s_1\mathbf{i}$, and so on, where

$$\mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \text{ and } \mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}.$$

(7) Verify directly that the extended distributive result $\mathbf{d}.(\mathbf{a}+\mathbf{b}+\mathbf{c}) = \mathbf{d}.\mathbf{a}+\mathbf{d}.\mathbf{b}+\mathbf{d}.\mathbf{c}$ is true for the case where $\mathbf{a} = p\mathbf{i}$, $\mathbf{b} = q\mathbf{j}$, $\mathbf{c} = r\mathbf{k}$. How could the extended result be proved in general, assuming it is true for $\mathbf{d}.(\mathbf{a}+\mathbf{b})$?

(8) Establish the following theorem:

If non-zero vectors $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ are orthogonal they are linearly independent.

(*Hint*: Suppose that they are dependent, i.e. λ and μ exist not both zero such that $\lambda\mathbf{a}+\mu\mathbf{b} = \mathbf{0}$: form the inner product of **a** with both sides. A contradiction will show your supposition to be false.)

(9) Expand $(\mathbf{a}+\mathbf{b}).(\mathbf{a}-\mathbf{b})$, using the distributive law. Hence show that if two vectors, which are neither equal nor opposite, have the same modulus, their sum and difference are orthogonal. What light does this throw on the properties of a rhombus?

(10) Express all the directed lines in the figure in terms of **p**, **q**, **h**. Given that $BH \perp AC$ and $CH \perp AB$, express these data in vector form and deduce that $AH \perp BC$.

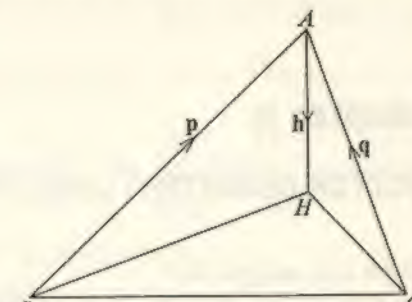


FIG. 28

(11) Prove the theorem of exercise 7a, question 6, using the result $|\mathbf{x}-\mathbf{a}|^2 = (\mathbf{x}-\mathbf{a}).(\mathbf{x}-\mathbf{a})$.

(12) A cubical box has $ABCD$ as one face and AA', BB', CC', DD' as parallel edges. X is the midpoint of the face $A'B'C'D'$, Y the midpoint of the edge $B'C'$ and Z the midpoint of the edge BC. Find the cosines of the angles $Y\hat{A}Z$, $Z\hat{A}X$, $X\hat{A}Y$. Express the volume of the tetrahedron $AXYZ$ as a fraction of the volume of the cube.

(13) The figure shows the two 3-vectors **a**, **b** with moduli a, b and direction-cosines (l_1, l_2, l_3), (m_1, m_2, m_3), the angle ϕ between them being acute. By projecting on to the line OB the vectors of the equation $\overline{OA} = \overline{OP}+\overline{PQ}+\overline{QA}$, establish the result $\cos\phi = l_1m_1+l_2m_2+l_3m_3$.

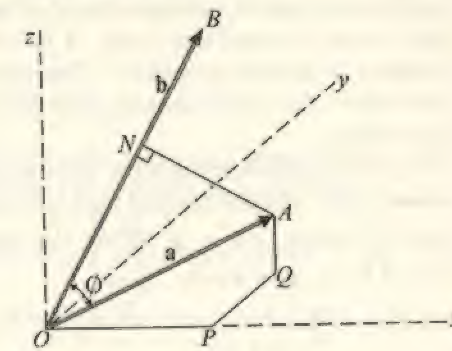


FIG. 29

State the geometrical theorem for projections which you would need in order to prove the result for any figure.

chapter 8

VECTOR GEOMETRY: LINES, PLANES, SURFACES

8.1 *What is a straight line?*

Since our procedure is to define our terms algebraically and to regard geometry as an embodiment of our algebraic ideas, we must be able to do this consistently from the beginning: we do this by regarding all the points in a given (geometrical) space as derived from our vectors such as $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ etc, with the operations and laws which these have.

It follows that we get a *set* of points if we give different values to the number t in the vector expression $t\mathbf{b}$. Such a set of points is called a *line* (formerly called a straight line).† Every value of t gives exactly one point on the line and vice versa. A number t which acts in this way in geometry is called a *parameter*. One particular value, $t = 0$, gives the zero vector: this shows that we only derive lines through the origin by this method.

To get a parametric form for *any* line, we only need to start at some chosen point A in the line and push out, as it were, from A to a variable point X, where $\overline{AX} = t\mathbf{b}$. Then the position-vector of X is $\overline{OX} = \overline{OA} + \overline{AX}$, i.e. $\mathbf{x} = \mathbf{a} + t\mathbf{b}$.

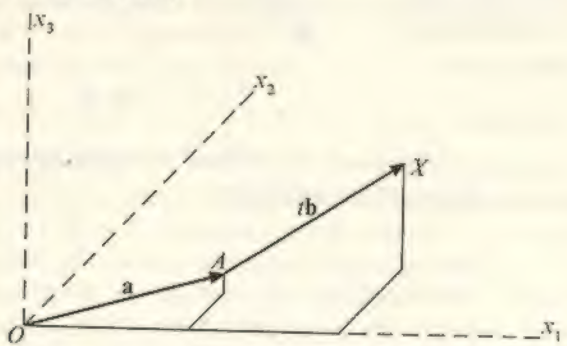


FIG. 30

† We distinguish a *ray* or *half-line* (e.g. for which $t \geqslant t_0$) from a *line-segment* (e.g. $t_1 \leqslant t \leqslant t_2$) and a *line* for which t takes all values.

The vector $\mathbf{b}$ gives the direction of the line, and can if desired be a unit-vector $\hat{\mathbf{l}} = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$ where l_1, l_2, l_3 are direction-cosines. Then $\mathbf{x} = \mathbf{a} + t\hat{\mathbf{l}}$, where t is now the *actual distance* AX; i.e. $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + t \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$

so that $x_1 = a_1 + tl_1$; $x_2 = a_2 + tl_2$; $x_3 = a_3 + tl_3$.

The third result is shown in Fig. 31 in the form $x_3 - a_3 = t \cos\theta_3$, since $AX = t$.

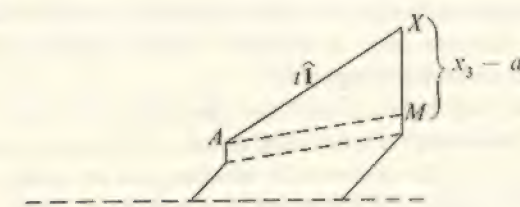


FIG. 31

Worked example

To find in the parametric form $\mathbf{x} = \mathbf{a} + t\mathbf{l}$, the line joining $A(1, -3, 2)$ to $C(4, 1, 14)$, and two points on it at distance $6\frac{1}{2}$ units from A.

Solution: The vector AC (or $\mathbf{c} - \mathbf{a}$) is $\begin{pmatrix} 3 \\ 4 \\ 12 \end{pmatrix}$ of modulus 13.

Thus $\overline{AX}$, the vector to any point on the line, has direction-numbers 3, 4, 12, and we can write $\mathbf{x} = \mathbf{a} + t\begin{pmatrix} 3 \\ 4 \\ 12 \end{pmatrix}$, or $\begin{matrix} x_1 = 1 + 3t \\ x_2 = -3 + 4t \\ x_3 = 2 + 12t. \end{matrix}$

If we wish the parameter to be actual distance we use the direction-cosine vector $\begin{pmatrix} 3/13 \\ 4/13 \\ 12/13 \end{pmatrix}$ and write $\mathbf{x} = \begin{pmatrix} 3 \\ -3 \\ 2 \end{pmatrix} + s\begin{pmatrix} 3/13 \\ 4/13 \\ 12/13 \end{pmatrix}$.

In full co-ordinate form this would be:

$$x_1 = 1 + \frac{3s}{13}; \quad x_2 = -3 + \frac{4s}{13}; \quad x_3 = 2 + \frac{12s}{13}.$$

Less usefully:

$$\frac{x_1-1}{3}=\frac{x_2+3}{4}=\frac{x_3-2}{12}\left(\text{each equal to } t \text{ or } \frac{s}{13}\right).$$

Points on the line distant $6\frac{1}{2}$ units from A will be given by writing $s = \pm 6\frac{1}{2}$, i.e. $(\frac{5}{2}, -1, 8)$ and $(-\frac{1}{2}, -5, -4)$.

EXERCISE 8a

(1) Find in form $\mathbf{x} = \mathbf{a}+t\mathbf{l}$ the equation of the line joining $A(-5, 1, 7)$ and $B(-1, 3, 3)$. Find a point P on the line distant 6 units from A on the side opposite from B.

(2) With the data of question 1, find the point Q which is 9 units from B on the side opposite from A; and also R, the midpoint of PQ. What is your t-value corresponding to R?

(3) Show that any point of AB (in question 1) can be written as $\mathbf{a}+t(\mathbf{b}-\mathbf{a})$. State the positions of the points which are given by writing $t = 0, \frac{1}{2}, 2, \frac{1}{3}$. Find the value of t which gives the point $\frac{1}{7}(4\mathbf{a}+3\mathbf{b})$ and explain the relation of this point to A and B.

(4) Points S, T are obtained by writing s, t for the parameter λ in the line $\mathbf{x} = \lambda\mathbf{a}$; points S', T' are obtained by writing s, t for μ in the line $\mathbf{x} = \mu\mathbf{b}$. Show that SS', TT' are parallel: compare their magnitudes, and comment on the result.

(5) In the triangle ABC, show that the median from A can be written $\mathbf{a}+\frac{1}{2}t\,(\mathbf{b}+\mathbf{c}-2\mathbf{a})$. Which special point is obtained by writing $t = 1$? Show that the point given by $t = \frac{2}{3}$ also lies on the other medians.

(6) Show that the lines joining $(6, 3, 4)$ to $(-3, 0, 1)$ and $(-7, -3, 8)$ to $(1, 1, 4)$ intersect in the point $(3, 2, 3)$.

(7) Find the foot of the perpendicular from the origin on to the line joining $(5, 3, 5)$ to $(8, 4, 7)$.

(8) Show that the perpendicular distance from the point $(9, -11, 17)$ to the join of $(7, 2, 4)$ and $(11, 3, 3)$ is 18 units.

8.2 *What is a plane?*

We shall consider first a plane through the origin. We have a clear geometrical picture of such a plane. By the vector approach, the points of the plane are given by their position-vectors, and these are derived from two base-vectors, e.g. $\mathbf{a}$, $\mathbf{b}$, every point being expressible as $\lambda\mathbf{a}+\mu\mathbf{b}$. (Euclid regarded a plane as defined by two distinct lines in it, and this is what we have here.)

An important property of a plane is that it has a unique *normal* at any point P: a line through P which is perpendicular to *every* line in the plane. We evidently have here a tremendous bonus, for we have no right to expect such a result. (We saw in section 7.3, worked example (*b*), how to find a direction perpendicular to any *two* given lines; but it appears now that when this has been found this direction has the property of being perpendicular to *all* the lines coplanar with the original two.)

In algebra this 'normal' property follows at once from the distributive law for our vectors:

Suppose we have a direction $\mathbf{l}$ orthogonal to the vectors $\mathbf{a}$ and $\mathbf{b}$: is it true that $\mathbf{l}$ is orthogonal to $\lambda\mathbf{a}+\mu\mathbf{b}$ for all λ, μ? i.e. is it true that $\mathbf{l}.\mathbf{a} = 0$ and $\mathbf{l}.\mathbf{b} = 0 \Rightarrow \mathbf{l}.(\lambda\mathbf{a}+\mu\mathbf{b}) = 0$? Clearly the answer is yes, because the expression equals $\lambda\mathbf{l}.\mathbf{a}+\mu\mathbf{l}.\mathbf{b}$.

This approach gives us a neat way of writing the vector equation of a plane through the origin: if the normal is written $\hat{\mathbf{n}}$ (a unit-vector) then a point with position-vector $\mathbf{x}$ will lie in the required plane provided that $\mathbf{x}.\hat{\mathbf{n}} = 0$, i.e. in co-ordinate form $x_1n_1+x_2n_2+x_3n_3 = 0$ (in which direction-cosines could be replaced by direction-numbers).

In a similar way a plane with normal $\hat{\mathbf{n}}$, passing through a given point A (with position-vector $\mathbf{a}$) is given by $(\mathbf{x}-\mathbf{a}).\hat{\mathbf{n}} = 0$.

This may also be written as $\mathbf{x}.\hat{\mathbf{n}} = \mathbf{a}.\hat{\mathbf{n}}$ or $\mathbf{x}.\hat{\mathbf{n}} = h$, where h is $\mathbf{a}.\hat{\mathbf{n}}$, the projection of OA in the direction $\hat{\mathbf{n}}$.

Notice that the equation $\mathbf{x}.\hat{\mathbf{n}} = h$ is satisfied by every vector $\mathbf{x}$ such that $OX \cos\phi$ (see diagram) is equal to h, which we identify as the length of the normal from O to the plane—counted positive if it is in the direction $\hat{\mathbf{n}}$, or negative if in the opposite direction.

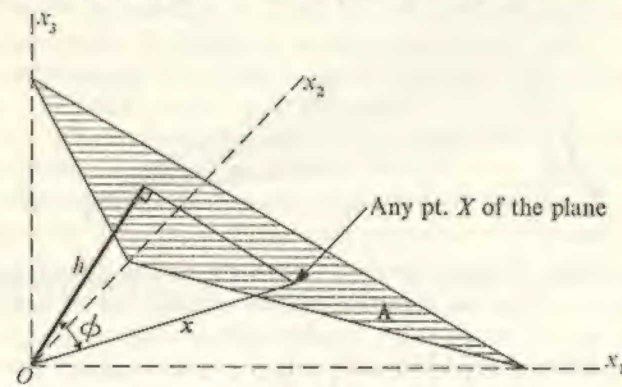


FIG. 32

Notice that our equation of a plane is not a parametric equation but a locus-equation—a condition to be satisfied by the appropriate values of $\mathbf{x}$.

Worked examples

(*a*) A plane is normal to the line with direction-numbers $1, 2, -2$, and the perpendicular to it from the origin is 6 units long (drawn in the sense given by the numbers $1, 2, -2$). Find where the plane meets the axes.

Solution: Since $1^2+2^2+2^2 = 9$, the direction-cosines are $\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}$ and the equation $\mathbf{x}.\hat{\mathbf{n}} = 6$.

If $\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$ satisfies, we have $an_1 = 6$, i.e. $a = 18$.

Similarly, $bn_2 = 6$ gives $b = 9$, and finally $c = -9$.

Notice that the trigonometry of this question gives us a new picture of the direction-cosines: the plane cuts off a, b, c from the axes, and the cosines are seen directly from the figure to be $n_1 = \cos\theta_1 = \frac{h}{a}$; $n_2 = \cos\theta_2 = \frac{h}{b}$; and $n_3 = \cos\theta_3 = \frac{h}{c}$.

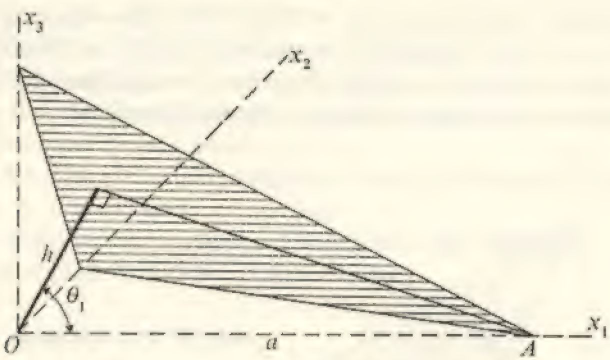


FIG. 33

(*b*) Two planes through the origin have normals in directions $1, 3, 2$, and $-1, 1, 2$. Find the direction-cosines of their line of intersection.

Solution: Write $\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$.

We have $\mathbf{x}.\mathbf{a} = 0$ i.e. $x_1+3x_2+2x_3 = 0$, and similarly $-x_1+x_2+2x_3 = 0 \Rightarrow 4x_2+4x_3 = 0$ i.e. $x_3 = -x_2$ and $x_1 = x_2+2x_3 = -x_2$.

Thus $x_1 : x_2 : x_3 = 1 : -1 : 1$; i.e. the direction-cosines are $\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$, or the same with reversed signs.

EXERCISE **8b**

(1) Show that the line joining $(2, 0, -3)$ to $(6, -2, 1)$ is bisected at right angles by the plane $2x-y+2z = 7$.

(2) Find the image of the origin formed by reflection in the plane $2x+6y+3z = 49$.

(3) Find the equation of the plane through $(1, 0, -1)$ at right angles to the planes $2x+5y+z = 0$ and $x-2y+2z = 0$.

(4) Find the equation of the line of intersection of the planes $x+y+z = 0$ and $3x-4y+5z = 5$.

(5) Show that the three planes $x+2y-2z = 0, 4x-y+z = 9, y+z = 3$ are mutually perpendicular, and find their point of intersection.

(6) The perpendicular from the origin on to a plane π is 13 units long and has direction-numbers (correct for sign) of 4, 12, 3 respectively. Write the equation of the plane in two forms, and find its intercepts on axes.

(7) (Continuing question 6): Find (*i*) the Cartesian equations, (*ii*) the direction-cosines of the line L_1 where π meets the plane x_1Ox_2. Hence find the direction of a line L_2 in π, which is perpendicular to L_1. (*Hint*: The line L_2 is also $\perp$ to the *normal*.)

(8) Two planes have normals in directions $2, -2, 1$ and $3, 5, 1$. Find direction-numbers for the line of intersection, and also a parametric equation for this line if it passes through the point $(-1, 4, 2)$.

(9) Show by two methods that the co-ordinate equation of a plane making intercepts a, b, c on the axis is $\frac{x}{a}+\frac{y}{b}+\frac{z}{c} = 1$.

A regular octahedron has vertices $(\pm a, 0, 0)$, $(0, \pm a, 0)$, $(0, 0, \pm a)$: find the direction-cosines of the outward normals to its faces, and the distance from the origin to a face.

8.3 *Various techniques with lines and planes*

In the previous section we considered a plane through the origin in some detail, in the forms:

(*a*) $\mathbf{x}.\hat{\mathbf{n}} = 0$ where $\hat{\mathbf{n}}$ is a unit-vector normal to the plane. Note that $-\hat{\mathbf{n}}$ can also be used. This is not parametric in form.

(*b*) $\mathbf{x} = p\mathbf{a}+q\mathbf{b}$ ($\mathbf{a}, \mathbf{b}$ linearly independent) which can be described as a two-parameter form, since every pair (p, q) of values of p and q gives a point on the plane and vice versa. Notice that in the case in which $\mathbf{a}, \mathbf{b}$ are $\mathbf{i}, \mathbf{j}$, we have $\mathbf{x} = p\mathbf{i}+q\mathbf{j} = \begin{pmatrix} p \\ q \\ 0 \end{pmatrix}$. This means that every point in the plane x_1Ox_2 is specified by its pair of co-ordinates $x_1 = p, x_2 = q$.

When considering a plane not through the origin we wrote $\mathbf{x}.\hat{\mathbf{n}} = h$; the reader will readily extend form (*b*) into $\mathbf{x} = p\mathbf{a}+q\mathbf{b}+\mathbf{c}$, where $\mathbf{c}$ is the position-vector of some point C on the plane.

The necessary and sufficient conditions for the plane defined by (*b*) to be the same as that defined by (*a*) are seen *from the geometry* to be:

(*i*) $\mathbf{a}, \hat{\mathbf{n}}$ orthogonal (*ii*) $\mathbf{b}, \hat{\mathbf{n}}$ orthogonal (*iii*) $\mathbf{c}.\hat{\mathbf{n}} = h$.

(*Hint*: Use (*iii*) first.)

The reader should verify these by means of a figure. Sufficiency means (*i*) (*ii*) and (*iii*) $\Rightarrow$ the planes are the same; while necessity means that 'the planes are the same' $\Rightarrow$ (*i*) (*ii*) and (*iii*).

We shall not consider here an algebraic treatment: it is enough that we should rather expect three linear ('length') conditions to arise where a point is specified by three 'lengths'.

Numerical examples on intersections of lines and planes are usually done most easily by a combination of vector and co-ordinate methods. The vectors give concise notation, but calculation must be done with components, i.e. co-ordinates.

Worked examples

(*a*) Find the point of intersection of the line PQ and the plane through A with normal given by $-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}$, where A is (0, 10, 23), P is (2, 0, 5) and Q is (5, 12, 9).

Solution: The plane is $\mathbf{x}.\hat{\mathbf{n}} = h$ where $\mathbf{a}.\hat{\mathbf{n}} = h$,

i.e. $-\frac{1}{3}x_1+\frac{2}{3}x_2+\frac{2}{3}x_3 = 0+\frac{2}{3}.10+\frac{2}{3}.23 = \frac{66}{3}$

or $-x_1+2x_2+2x_3 = 66.$

The vector $\overline{PQ}$ is $\mathbf{q}-\mathbf{p} = \begin{pmatrix} 3 \\ 12 \\ 4 \end{pmatrix}$ and the line PQ is $\mathbf{p}+t(\mathbf{q}-\mathbf{p}) =$

$\begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} + t\begin{pmatrix} 3 \\ 12 \\ 4 \end{pmatrix}$; i.e. for the required point $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2+3t \\ 12t \\ 5+4t \end{pmatrix}$.

Then, by substitution:

$$-(2+3t)+24t+10+8t = 66, \text{ i.e. } 29t = 58 \Rightarrow t = 2,$$

and the point required is (8, 24, 13).

(*b*) A line through the origin with direction-numbers 0, 3, 5 meets the plane through P(6, 1, 14) which has normal in direction 1, 1, −2. Find the point of intersection.

Solution: We will first do it by a general vector method.

Write $\mathbf{a} = \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}$, so that $\mathbf{x} = t\mathbf{a}$ is the line. Write $\mathbf{p} = \begin{pmatrix} 6 \\ 1 \\ 14 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$, so that the plane $\mathbf{x}.\mathbf{b} = h$ where h is such that p lies on the plane; i.e. $\mathbf{x}.\mathbf{b} = \mathbf{p}.\mathbf{b}$. (*Note*: $\mathbf{b}$ has not here been made into a unit-vector.)

Then, given $\mathbf{x} = t\mathbf{a}$ and $\mathbf{x}.\mathbf{b} = \mathbf{p}.\mathbf{b}$, we are required to find the value of t. With only *one* piece of information required we can afford to substitute.†

Thus, $t\mathbf{a}.\mathbf{b} = \mathbf{p}.\mathbf{b}$, which gives t:

$$t(0+3-10) = 6+1-28, \text{ i.e. } -7t = -21 \Leftrightarrow t = 3$$

and this gives the point as (0, 9, 15).

Alternatively, using numbers and co-ordinates from the start:

$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3t \\ 5t \end{pmatrix}$ is to satisfy a plane of the form $x_1+x_2-2x_3 = k$, where k is such that 6, 1, 14 satisfies it; i.e. $k = -21$.

Then $0+3t-10t = -21 \Leftrightarrow t = 3$, and the point is (0, 9, 15).

(*c*) Two planes pass through the point A(5, 0, −1) and have normals given by direction numbers 2, 1, 3 and −1, 0, 1. Find the direction of the line of intersection and express the line in parametric form, with $t = 0$ to represent A.

† A scalar equation gives only one piece of information.

Solution: To find the direction required, we can transfer to parallel planes through the origin, i.e.:

$$\begin{aligned} 2x+y+3z &= 0 \\ -x \quad + z &= 0. \end{aligned}$$

For all the common points, $x = z$ and $y = -2x-3z = -5x$; i.e. $x : y : z = 1 : -5 : 1$, which are the direction numbers required for the line of intersection. They also apply to the original line, which can therefore be written as $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ -1 \end{pmatrix} + t\begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix}$, i.e. $x = 5+t$; $y = -5t$; $z = -1+t$.

EXERCISE 8c

(1) Of the two pairs of lines given below, one pair intersects while the other does not. Investigate which is which and find the point of intersection.

(*a*) $\mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + \lambda\begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} + \mu\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

(*b*) $\mathbf{x} = \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} + s\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} 3 \\ -2 \\ -5 \end{pmatrix} + t\begin{pmatrix} 5 \\ 3 \\ 7 \end{pmatrix}$

(2) Find the equation of the plane, normal to the vector with direction numbers 3, 4, 5, passing through the point (1, 1, 1). What would be the plane normal to the vector with direction a, b, c, through the same point?

(3) A plane is defined by the vectors $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$ and the point $(2, -1, 3)$. Find where the line $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + s\begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix}$ meets this plane.

(4) Two planes $x+3y-z = 4$ and $2x-y+z = 10$ meet in a line. Find the direction cosines of the line.

(5) The diagram shows a section of a surface in the neighbourhood of

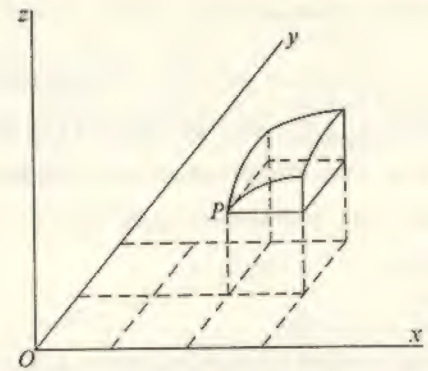


FIG. 34

a point $P(2, 1, 4)$. The tangent plane intersects the plane $y = 1$ in a line of gradient $\frac{1}{2}$, and the plane $x = 2$ in a line of gradient 1. Find the equation of the plane. This figure gives an interpretation for $\dfrac{\partial z}{\partial x}, \dfrac{\partial z}{\partial y}$ at the point P.

(6) The vertices of a tetrahedron are $A(0, 3, 2)$, $B(4, 0, 1)$, $C(-1, 1, 3)$ and $D(2, -1, 0)$. Find the direction-numbers of the vector perpendicular to the plane BCD, and hence write down in vector form the equation of the perpendicular from A to BCD. Do the same for B and plane ACD. Do these perpendiculars meet?

(7) Find the equation of the plane through the point (2, 1, 3) perpendicular to the line joining (2, 1, 3) and (4, 0, 5). What is the distance of (3, 1, 5) from this plane? Is this point on the same side of the plane as the origin?

(8) Show that $\mathbf{x} = p\mathbf{a}+q\mathbf{b}+r\mathbf{c}$ is a plane through the points with position-vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ iff† $p+q+r = 1$ is satisfied by the chosen values p, q, r. (*Hint*: $\mathbf{x}-\mathbf{a}, \mathbf{x}-\mathbf{b}, \mathbf{x}-\mathbf{c}$ are coplanar if a condition is satisfied which can be written as $\lambda(\mathbf{x}-\mathbf{a})+\mu(\mathbf{x}-\mathbf{b})+\nu(\mathbf{x}-\mathbf{c}) = \mathbf{0}$.)

(9) Consider some results in plane co-ordinate geometry of the straight line:

(*i*) parametric form $x = a+t\cos\theta$, $y = b+t\sin\theta$;

(*ii*) the 'α, p' form, where α is the slope of the normal to the line, viz. $x\cos\alpha+y\sin\alpha = p$.

† A short notation for 'if and only if'. The reader is to prove the implication both ways.

The reader should look closely at these forms in the light of the vector work done in sections 8.1-8.3 and discuss what takes the place of our 3-dimensional direction cosines when we work in two dimensions only.

(10) Restate the locus equation $ax+by = c$ in a form involving a scalar product. The vector $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{r}$ may be taken as the position-vector of a point on the locus: what interpretation can be given to (*i*) the vector $\begin{pmatrix} a \\ b \end{pmatrix}$; (*ii*) the vector with components $a/\sqrt{a^2+b^2}$, $b/\sqrt{a^2+b^2}$; (*iii*) the value of $c/\sqrt{a^2+b^2}$?

8.4 *Two lines in space*

Let A, B be two points given by position-vectors $\mathbf{a}$, $\mathbf{b}$, and let $\mathbf{x} = \mathbf{a}+s\mathbf{l}$ and $\mathbf{x} = \mathbf{b}+t\mathbf{m}$ be two non-parallel lines in space. In general they will be 'skew lines', i.e. they will not meet: the condition that they should do so is that s and t exist such that $\mathbf{a}-\mathbf{b}+s\mathbf{l}-t\mathbf{m} = 0$ . . . (E)

We can interpret this result thus: if we draw directions $\mathbf{l}$ and $\mathbf{m}$ through any chosen point in space, e.g. the origin, then $s\mathbf{l}-t\mathbf{m}$ is a direction in the plane containing $\mathbf{l}$ and $\mathbf{m}$. In fact, if we let s and t have all values we shall get this whole plane π-parallel to the two given lines. Equation (E) then expresses the fact that the join of A to B is parallel to this plane.

If, on the other hand, there is *no* meeting of the lines, then $\mathbf{a}-\mathbf{b}$ is *not* parallel to π and AB will meet π and any plane parallel to it. In particular we can draw a π-plane through the midpoint K of AB.

The situation at this stage is shown in figure 35. It is clear that we can *always* draw such a plane, given A, B, $\mathbf{l}$, $\mathbf{m}$, even if the lines AP, BQ actually meet, but it is when they are skew that this plane is particularly useful: in fact, to deal with skew lines one often contrives to take this as the co-ordinate plane x_1Ox_2.

A surprising result is that we get the same plane whichever pair of points we take on AP, BQ. As we defined it we obtained $\mathbf{k}+\lambda\mathbf{l}+\mu\mathbf{m}$, where λ, μ have any values; i.e. $\frac{1}{2}(\mathbf{a}+\mathbf{b})+\lambda\mathbf{l}+\mu\mathbf{m}$, a plane given in two-parameter form.

If, however, we had taken points $\mathbf{a}+s\mathbf{l}$, $\mathbf{b}+t\mathbf{m}$, we should have obtained $\frac{1}{2}(\mathbf{a}+\mathbf{b}+s\mathbf{l}+t\mathbf{m})+\lambda'\mathbf{l}+\mu'\mathbf{m}$, which is the same plane because the point obtained by putting λ, μ in the first is the same as the point on the second for which $\lambda'+\frac{1}{2}s = \lambda$ and $\mu'+\frac{1}{2}t = \mu$.

The plane could very well be called the mediant plane of the skew lines.

If we visualise the mediant plane π as horizontal, then AP is a horizontal line above π and BQ a horizontal line below it. By projecting down (and up) we can obtain two lines *in* the plane π and in general these will meet (i.e. always, unless $\mathbf{l}$, $\mathbf{m}$ are themselves parallel, which we assumed at the start not to be so.)

The new figure which emerges reveals that we can find in this way a cross line RS from AP to BQ which is perpendicular to π and consequently to both lines.

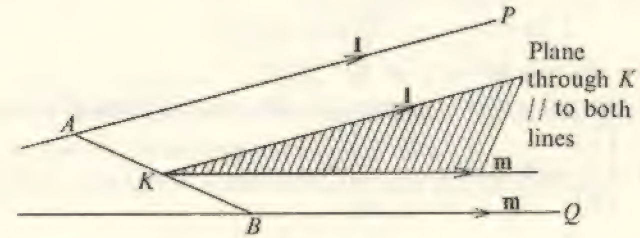


FIG. 35

If we take co-ordinate axes through O, the midpoint of RS, with Ox_1, Ox_2 bisecting the angle of $\mathbf{l}$, $\mathbf{m}$, and Ox_3 along OR, we can write

RP parametrically as $\begin{pmatrix} s\cos\alpha \\ s\sin\alpha \\ h \end{pmatrix}$ and SQ as $\begin{pmatrix} t\cos\alpha \\ -t\sin\alpha \\ -h \end{pmatrix}$ where 2α is the angle between $\mathbf{l}$, $\mathbf{m}$ and $2h$ is the length of RS.

We now verify that every midpoint of a crossline (e.g. of PQ) lies in x_1Ox_2, its co-ordinates being $(\overline{s+t}\cos\alpha, \overline{s-t}\sin\alpha, 0)$.

It should be noted that if $\mathbf{l}$ and $\mathbf{m}$ are themselves orthogonal there is nothing gained by using the bisectors as axes: we still use the mediant plane, but take Ox_1, Ox_2 parallel to $\mathbf{l}$, $\mathbf{m}$ so that the parametric forms are $\begin{pmatrix} s \\ 0 \\ h \end{pmatrix}\begin{pmatrix} 0 \\ t \\ -h \end{pmatrix}$.

Worked examples

(*a*) Prove that if a tetrahedron has two pairs of opposite sides perpendicular then the third pair are perpendicular. Take the vertices to be S, T, U, V, where ST, UV are $\perp$ and also SU, TV. Take the mediant

plane of the first pair as x_1Ox_2 where Ox_3 is the unique line perpendicular to both. Then we can write the points S, T, U, V as

$$\begin{pmatrix} s \\ 0 \\ h \end{pmatrix} \begin{pmatrix} t \\ 0 \\ h \end{pmatrix} \begin{pmatrix} 0 \\ u \\ -h \end{pmatrix} \begin{pmatrix} 0 \\ v \\ -h \end{pmatrix}.$$

Then $SU \perp TV \Leftrightarrow st+uv+4h^2 = 0$.

Also, SV is $\begin{pmatrix} s \\ -v \\ 2h \end{pmatrix}$ and TU is $\begin{pmatrix} t \\ -u \\ 2h \end{pmatrix}$,

therefore $st+uv+4h^2 = 0 \Leftrightarrow SV \perp TU$.

(*b*) Find the parametric equation of the line of intersection of the planes $\mathbf{x}.\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = 5$ and $\mathbf{x}.\begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} = 1$.

Take parallel planes through O, viz. $\begin{cases} x_1+2x_2+2x_3 = 0 \\ -2x_1+2x_2+x_3 = 0. \end{cases}$

Then for a point of intersection of these we have $3x_1+x_3 = 0$. Thus if $x_1 = 1, x_3 = -3$ and by substitution $x_2 = \frac{5}{2}$, i.e. cleared of fractions, 2, 5, -6 are solutions, and the line is

$\mathbf{x} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + t\begin{pmatrix} 2 \\ 5 \\ -6 \end{pmatrix}$ where $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ is any point on the two given planes.

Taking $a_3 = 0$, we have $\left.\begin{aligned} a_1+2a_2 &= 5 \\ -2a_1+2a_2 &= 1 \end{aligned}\right\}$

$$\Leftrightarrow a_1 = \tfrac{4}{3},\ a_2 = \tfrac{11}{6}.$$

Therefore the line is $\mathbf{x} = \begin{pmatrix} 4/3 \\ 11/6 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ 5 \\ -6 \end{pmatrix} = \begin{pmatrix} 4/3+2t \\ 11/6+5t \\ -6t \end{pmatrix}$.

The reader may notice that writing $a_3 = 1$ gives $a_1 = a_2 = 1$ but one could easily miss this.

If we write $t = -\frac{1}{6}$ in the result given, we get $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

and the simplest form of answer is $\begin{pmatrix} 1+2s \\ 1+5s \\ 1-6s \end{pmatrix}$.

EXERCISE **8d**

(1) Show that the lines $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + s\begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix}$ intersect. Give the parametric equation of the common normal at the point of intersection.

(2) Find the least distance apart of the lines

$$\mathbf{x} = t\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + s\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

(*Hint*: If P, Q are any points on the respective lines, find their parameters so that $\overline{PQ}$ is orthogonal to both lines.)

(3) Find the parametric equation of the line of intersection of

$$\mathbf{x}.\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \text{ and } \mathbf{x}.\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = 1.$$

(4) The plane containing $\mathbf{a}$, $\mathbf{b}$ has a normal in direction $\mathbf{n}$. Prove that

$$n_1:n_2:n_3 = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} : \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} : \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

(5) Given $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ the position vectors of A, B, C, form the position vectors of (*i*) M, the midpoint of BC; (*ii*) K in AM such that $AK = \frac{2}{3}AM$. Hence show that the medians of the $\triangle ABC$ concur. Prove similarly that in any tetrahedron $ABCD$ the lines joining vertices to the centroids of the opposite faces concur.

(6) A plane has a perpendicular distance p from the origin. The normal from O towards the plane has direction cosines l_1, l_2, l_3. The intercepts cut off from Ox_1, Ox_2, Ox_3 by the plane are a, b, c (with due regard to sign). Show that $p = al_1 = bl_2 = cl_3$ and hence or otherwise that the equation of the plane is $\frac{x}{a}+\frac{y}{b}+\frac{z}{c} = 1$.

(7) (*i*) Show that the line $\mathbf{x} = \begin{pmatrix} -4 \\ -3 \\ 1 \end{pmatrix} + t\begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix}$ can be written in the form $\frac{x+4}{3} = \frac{y+3}{2} = \frac{z-1}{-2}$. Hence or otherwise show that:

(*ii*) two planes containing the line are: $2x-3y-1 = 0$

$$y+z+2 = 0;$$

(*iii*) *any* plane containing the line can be written in the form $k(2x-3y-1)+k'(y+z+2) = 0$, where the ratio k to k' is suitably chosen.

(8) (Continuing question 7) The line L_1 has equation

$$\frac{x+4}{3} = \frac{y+3}{2} = \frac{z-1}{-2}.$$

Find the equation of a plane which contains L_1 and is perpendicular to the plane π with equation $x+y+z = 0$. Hence show that the projection of the line L_1 on plane is the line $\frac{x}{2} = y = \frac{z}{-3}$.

8.5 *Vector equations of some surfaces*

The sphere: Clearly $|\mathbf{x}| = R$ is the equation of a sphere with centre O and radius R. It can also be written $x_1^2+x_2^2+x_3^2 = R^2$.

The points on the sphere for which $x_3 = h$ are given by

$$\left.\begin{aligned} x_1^2+x_2^2 &= R^2-h^2 \\ x_3 &= h \end{aligned}\right\} \text{(a circle)}$$

If we were to omit the second condition, x_3 would be unlimited and we should have a circular cylinder with axis Ox_3.

A sphere of centre A would be $|\mathbf{x}-\mathbf{a}| = R$

or $\qquad (x_1-a_1)^2+(x_2-a_2)^2+(x_3-a_3)^2 = R^2.$

A simple (i.e. right circular) cone with vertex at the origin

The vector $\mathbf{x}$ is required to make a fixed angle with a given line $\mathbf{l}$ (the axis of the cone), thus:

$$\mathbf{x}.\mathbf{l} = |\mathbf{x}||\mathbf{l}|\cos\alpha; \text{ or } \mathbf{x}.\hat{\mathbf{l}} = |\mathbf{x}|\cos\alpha \text{ if } \hat{\mathbf{l}} \text{ is a unit vector.}$$

In Cartesian form, by squaring we have:

$$(x_1l_1+x_2l_2+x_3l_3)^2 = (x_1^2+x_2^2+x_3^2)\cos^2\alpha.$$

This is a homogeneous equation, and if (x_1, x_2, x_3) is a point on it, so is px_1, px_2, px_3, a fact which is equally well shown by the vector equation.

EXERCISE 8e

(1) Find the radii of the circles in which the sphere $|\mathbf{x}| = R$ cuts:

(*i*) $\mathbf{x}.\hat{\mathbf{l}} = p$ where $\hat{\mathbf{l}}$ is a unit-vector and $p < R$;

(*ii*) $\mathbf{x}.\hat{\mathbf{l}} = |\mathbf{x}|\cos\alpha$.

(2) Describe the loci given in terms of a parameter t by

$$(i)\ \begin{pmatrix} a\cos t \\ a\sin t \\ h \end{pmatrix}; \qquad (ii)\ \begin{pmatrix} a\cos t \\ b\sin t \\ h \end{pmatrix}; \qquad (iii)\ \begin{pmatrix} a\cos t \\ a\sin t \\ ht \end{pmatrix}.$$

In each case express $\mathbf{x}$ in terms of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and t; and also express $\frac{d\mathbf{x}}{dt}$ in similar form. Show that for one of these loci the vectors $\mathbf{x}$ and $\frac{d\mathbf{x}}{dt}$ are always orthogonal.

By definition $\qquad \frac{d}{dt}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \frac{dv_1}{dt}.\mathbf{i}+\frac{dv_2}{dt}.\mathbf{j}+\frac{dv_3}{dt}.\mathbf{k}$

(3) A beacon B on the top of a cliff of height h above the sea projects light in all directions within a cone which includes in its surface the lines Bx_1, Bx_2, Bx_3, running due east, due north, and vertically upwards from B. If the direction of any ray from the beacon is given by unit-vector $\hat{\mathbf{l}}$, obtain an inequality governing its components, l_1, l_2, l_3.

Show that points at sea level from which light is just able to be received lie on one branch of a rectangular hyperbola, and find the nearest such point, measured from the cliff foot below B.

(4) X is a point such that $\overline{OX}$ makes an angle of 45° with the axis Ox_1. Give an equation which is satisfied by $\mathbf{x}$, the position-vector of X. (*Hint*: Use unit-vector $\mathbf{i}$.) Give the locus of X in Cartesian form. Show that if (x_1, x_2, x_3) satisfies the equation so also does (kx_1, kx_2, kx_3). In what curve does the locus cut the plane $x_2 = 1$?

(5) V is the point with position-vector $\begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix}$ and $\mathbf{l} = \begin{pmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{pmatrix}$. The point X is such that $\overline{VX}$ makes an angle θ with $\mathbf{l}$, where $\cos\theta = \dfrac{2}{\sqrt{5}}$. Show by a geometrical argument that X lies on the surface of a cone. Obtain the locus (*i*) in the form $(\mathbf{x}-\mathbf{v}).\mathbf{l} = k\,|\,\mathbf{x}-\mathbf{v}\,|$ where k is a numerical constant; (*ii*) in Cartesian form.

Hence show that the cone intersects the plane $x_1 = 4$ in points which satisfy the equation $x_2^2 = 4x_3$.

(6) Explain why the following implication is irreversible:

$$\mathbf{x}.\mathbf{l} = |\,\mathbf{x}\,|\cos\alpha \Rightarrow (x_1l_1+x_2l_2+x_3l_3)^2 = (x_1^2+x_2^2+x_3^2)\cos^2\alpha.$$

Interpret this fact in terms of the loci which the equations represent.

8.6 *Equivalence classes at work: a logical appendix*

Vector algebra is a complete and consistent system. When we apply it to geometry we have to appreciate, as has been discussed in chapter 2, that **a** is a symbol for a whole class of equal and parallel displacements, of which we may use only one or two in our geometrical figure: but if this is the situation, then an equation such as $\mathbf{a}+\mathbf{b} = \mathbf{c}$ represents a highly complicated situation in geometry. We are forced to regard it as representing *every* directed line of type **a**, each with a suitable line **b** added to it.† In order to write a unique class-symbol **c** for the result we must know (i.e. in fact *postulate*, if we are building up our geometry on the algebra) that the results are all equal and parallel and conversely that all the **c**-lines can be formed in this way; in short, that addition by classes yields a class. This type of class operation is known in other situations in algebra itself and is called *compatibility of classes under an extended operation*. We shall consider some examples, and (as always) a counter-example, to make this clear.

We discussed in chapter 2 the case of equivalence classes of integers to a modulus. A very easy example to handle is modulo 10, which regards as equivalent all denary numbers with the same final integer, e.g. 3, 13, 23, ...

† This shows that there is still a logical loophole, requiring a deeper treatment to dispose of it: addition of directed line-segments is not here *well-defined*, i.e. we are not able to add *any* pair.

It is readily seen that such classes are compatible under extended multiplication, e.g. denoting classes by C_1, C_2, C_3 etc.

$$C_2 \times C_3 = \begin{pmatrix} 2\times\ 3,\ 12\times\ 3,\ 22\times\ 3, \ldots \\ 2\times 13,\ 12\times 13,\ 22\times 13, \ldots \\ 2\times 23, \ldots \\ \ldots \end{pmatrix} = C_6$$

Remember that the basis operation stated is for the *members* of a class. We shall now show a counter-example. The operation $*$ applied to two numbers consists of adding the digits of each, and multiplying the results together; for example, $11 * 13 = 8$.

If applied to the modular classes it gives discrepant results, e.g. $2 * 3 = 6$ and $12 * 3 = 9$, so that it is not true to write $C_2 * C_3 =$ some class C_p.

A final exercise

The reader should verify that on the other hand the operation $*$, when carried out on integers modulo 9, is found to be compatible with the equivalence classes C_0, C_1, etc., to C_8. This method ('casting out the nines') has in fact been used for more than 300 years for testing† accuracy of multiplication. For example, 243 is of class C_0, and when multiplied by anything it should give a class C_0 number, since $C_0 * C_n = C_0$.

It is worth a comment in conclusion that the idea of an equivalence class underlies many situations, notably geometrical proofs of Euclidean style based on a figure. The figure is assumed to be a fair representative of all the figures which could be drawn. Whenever, in any context, we use a particular instance as a basis for a general argument we are assuming that all possible cases are like the example in all the respects which matter, i.e. it is claimed that an equivalence exists.

EXERCISE 8f (*Miscellaneous*)

(1) The vertices of a tetrahedron are $A(0, 0, 0)$, $B(4, 3, 0)$, $C(0, 5, 0)$, $D(4, 2, 4)$. Find (*i*) the angle between the lines AC, AD; (*ii*) the angle between the planes ABC, ABD; (*iii*) the volume of the tetrahedron.

(2) Find the equation of the plane which passes through the point (1, 1, 1) and contains the line of intersection of the planes

$$3x-y-z-2 = 0 \text{ and } 4x+y-z-5 = 0.$$

(*Hint*: If the planes are $P = 0$, $P' = 0$, what does $P+\lambda P' = 0$ represent?)

† It is a test, not a *check*. Test fails $\Rightarrow$ product false, but test succeeds $\not\Rightarrow$ product correct.

(3) Find the equation of the plane which contains the line $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ and is perpendicular to the plane $3x+2y-z = 0$.

(4) Find the equation of the line through $(-1, 5, 4)$ which meets the line $\mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} + t\begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$ and is parallel to the plane $x+2y-3z = 4$.

(5) An extensive cloud layer has a plane base. Taking rectangular axes Ox, Oy on the ground (supposed flat) and Oz upwards, the cloud-base meets the ground in the line $\dfrac{x}{150}+\dfrac{y}{200} = 1, z = 0$, units being miles; and a point 6 miles vertically above the origin is also in the plane. What is the inclination of the plane to the ground, and what is its height above the point $(50, 100, 0)$?

(6) Show that the equation of the line through A perpendicular to the plane $\mathbf{x}.\hat{\mathbf{n}} = h$ is $\mathbf{x} = \mathbf{a}+t\hat{\mathbf{n}}$ where $\mathbf{a}$ is the position-vector of A. Show also that the point with parameter $t = 2h-2\mathbf{a}.\hat{\mathbf{n}}$ is the reflected image of A in the plane.

(7) Show that there is no finite solution to the equations

$$\begin{aligned} 2x_1+5x_2+4x_3 &= 8, \\ 3x_1+5x_2+\ x_3 &= 3, \\ 2x_1+3x_2 \qquad &= 4, \end{aligned}$$

and interpret this result geometrically.

(8) Show that the homogeneous equation $x^2+y^2+z^2 = (l_1x+l_2y+l_3z)^2$ represents a cone, where l_1, l_2, l_3 are the components of a unit-vector. Interpret the cone in relation to the plane $l_1x+l_2y+l_3z = p$.

(9) Interpret the parametric equations:

(*i*)
$$\mathbf{x} = \begin{pmatrix} 30t \\ 0 \\ 40t-16t^2 \end{pmatrix}$$

(*ii*)
$$\mathbf{x} = \begin{pmatrix} e^t \cos t \\ e^t \sin t \\ 0 \end{pmatrix}$$

(*iii*)
$$\mathbf{x} = \begin{pmatrix} \cos t \cos 2t \\ \cos t \sin 2t \\ \sin t \end{pmatrix}$$

(10) The point A is $(0, 0, 2)$, B is $(-4, 2, 6)$, C is $(6, 3, 4)$ and D is $(-1, -4, -3)$. Show that the points are coplanar, and find whether AB intersects CD internally or externally. Find also the equation of the plane containing the points. (*Hint*: Form any plane through AB and adjust it to pass through C.)

(11) Find the equation of the line joining the origin to the point $(1, -1, 2)$. Hence or otherwise find the equation of the plane containing the origin and the line $\dfrac{x-1}{3} = \dfrac{y+1}{6} = \dfrac{z-2}{-2}$.

(12) Two lines OA, OB are drawn from the origin in the directions of the vectors $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$. Prove that all lines through O which are equally inclined to $\mathbf{a}$, $\mathbf{b}$ lie in the plane $x+5y-4z = 0$.

(13) Find the equation of the plane containing the point $(0, 2, -1)$ and the straight line $\dfrac{x+1}{2} = \dfrac{y}{2} = 3-z$.

(14) Find the foot of the perpendicular from the origin to the line $x = 7+t$, $y = -1-4t$, $z = 5+3t$.

(15) Find the angle between the planes $6x-3y+2z = 9$, $2x+2y-z = 0$, and the equation of their line of intersection.

(16) Solve:

(*a*)
$$\begin{aligned} x-3y+\ z &= 0 \\ 2x+\ y-4z &= 0 \end{aligned}$$

(*b*)
$$\begin{aligned} x-3y+\ z &= 0 \\ 2x+\ y-4z &= 0 \\ x+\ y-\ z &= 2 \end{aligned}$$

(*c*)
$$\begin{aligned} x-3y+\ z &= 8 \\ 2x+\ y-4z &= 4 \end{aligned}$$

(*d*)
$$\begin{aligned} x-3y+\ z &= 8 \\ 2x+\ y-4z &= 4 \\ x+\ y-\ z &= 24 \end{aligned}$$

(*e*)
$$\begin{aligned} x-3y+\ z &= 1 \\ 2x+\ y-4z &= 1 \\ 5x-8y-\ z &= 0 \end{aligned}$$

(*f*)
$$\begin{aligned} x-3y+\ z &= 1 \\ 2x+\ y-4z &= 1 \\ 5x-8y-\ z &= 4 \end{aligned}$$

(17) It is a common task in numerical work to fit a polynomial $p(x)$ to given data, e.g. to find a cubic polynomial such that $p(1) = a$, $p(2) = b$, $p(3) = c$, $p(4) = d$. Since cubic polynomials form a vector space of dimension four we solve by forming a suitable basis: the first base vector is the cubic $f_1(x)$ such that $f_1(1) = 1$ and f_1 vanishes for the other three values; the second vector is $f_2(x)$ such that $f_2(2) = 1$ but f_2 vanishes for $x = 1$, 3 and 4; and so on.

(*i*) Show that $f_1(x)$ is for the form $k_1(x-2)(x-3)(x-4)$ and calculate the required value of k_1.

(*ii*) Show that $f_1(x), f_2(x), f_3(x), f_4(x)$ are linearly independent.

(*iii*) Express the required polynomial as a linear combination of the base polynomials f_1, f_2, f_3, f_4.

(18) (*a*) All the vectors of a certain space V are expressible as $p\mathbf{i}+q\mathbf{j}$ where $\mathbf{i}, \mathbf{j}$ are unit-vectors and p, q are real numbers. A subset R of V consists of all the unit-vectors in V. Explain why R is not itself a vector space.

(*b*) For any vector $\mathbf{r}$ of R we define a real number $n(\mathbf{r})$ such that (*i*) $n(\mathbf{r})$ is numerically equal to $1-\mathbf{r}.\mathbf{i}$; and (*ii*) the sign of $n(\mathbf{r})$ is that of $\mathbf{r}.\mathbf{j}$.

Discuss this number n as an indicator of 'direction' in the 'plane'.

(*c*) For any pair $\mathbf{r}, \mathbf{s}$ of vectors in R we define a real number $m(\mathbf{r}, \mathbf{s})$ as an indicator of 'opening between directions' by the expression $m(\mathbf{r}, \mathbf{s}) = 1-\mathbf{r}.\mathbf{s}$. Discuss this number. If it is to be a *measure* it must certainly satisfy the conditions:

(*i*) $m(\mathbf{r}, \mathbf{r}) = 0$

(*ii*) $m(\mathbf{r}, \mathbf{s})+m(\mathbf{s}, \mathbf{t}) = m(\mathbf{r}, \mathbf{t})$, where addition could perhaps be *modular*, since angles can be added modulo 2π.

(19) The vector product of two 3-vectors $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ is defined by $\mathbf{a}\times\mathbf{b} = \begin{pmatrix} a_2b_3-a_3b_2 \\ a_3b_1-a_1b_3 \\ a_1b_2-a_2b_1 \end{pmatrix}$.

Show:

(*i*) that $\mathbf{a}\times\mathbf{b} \neq \mathbf{b}\times\mathbf{a}$;

(*ii*) that $\mathbf{a}\times\mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$, with a right-hand screw relation between $\mathbf{a}, \mathbf{b}, \mathbf{a}\times\mathbf{b}$ in that order;

(*iii*) that $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta$ where θ is the angle between the directions of $\mathbf{a}, \mathbf{b}$;

(*iv*) that 3-vectors do not form a group under this law of combination;

(*v*) that $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}.\mathbf{c})\mathbf{b}-(\mathbf{a}.\mathbf{b})\mathbf{c}$.

(*Note*: $\mathbf{i}\times\mathbf{j} = \mathbf{k}$ and a number of similar results; also the definition suggests that vector product is distributive over addition, e.g.

$$(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k})$$

can be multiplied out. These facts may be found useful.)

PART II

chapter 9

THE MACHINERY OF SIMPLE MATRICES

9.1 We assume that the reader is familiar, at least in simple cases, with the notation of matrices. In the first instance a matrix may be regarded as a rectangular array of numbers, a pattern of information.

Examples of matrices are:

$$\begin{pmatrix} 3 & 0 \\ 1 & 2 \\ 0 & 5 \end{pmatrix} \text{ and } \begin{pmatrix} 15 & -3 \\ 13 & -2 \end{pmatrix}.$$

These could have arisen in a commercial situation as follows:

$$\begin{array}{l} \\ \\ \text{Tin of type } \mathbf{A} \\ \text{Tin of type } \mathbf{B} \\ \text{Tin of type } \mathbf{C} \end{array} \begin{array}{c} \textit{Stock matrix} \\ \begin{array}{cc} \textit{Large} & \textit{Small} \end{array} \\ \begin{pmatrix} 3 & 0 \\ 1 & 2 \\ 0 & 5 \end{pmatrix} \end{array} \qquad \begin{array}{c} \textit{Cost Matrix} \\ \begin{array}{cc} \textit{Present cost} & \textit{Proposed change} \end{array} \\ \begin{array}{l} \text{large} \\ \text{small} \end{array} \begin{pmatrix} 15 & -3 \\ 13 & -2 \end{pmatrix} \end{array}$$

For our present purposes we can regard the stock matrix as a series of rows of information, and the cost matrix as a series of columns of information (row-vectors and column-vectors if preferred): these vectors necessarily contain the same number of items, since these items are in both cases keyed to the words 'large', 'small'.

We now form the inner product of the rows of the stock matrix with columns of the cost matrix, and find that each such product gives us an intelligible result, viz. a valuation (or change in valuation) of a section of the stock, thus:

	Present valuation
Type **A**	$3\times15 + 0\times13 = 45$
Type **B**	$1\times15 + 2\times13 = 41$
Type **C**	$0\times15 + 5\times13 = 65$

In fact, we could set out the results in a matrix:

$$\begin{array}{l} \\ \\ \text{Type } \mathbf{A} \\ \text{Type } \mathbf{B} \\ \text{Type } \mathbf{C} \end{array} \begin{array}{c} \textit{Valuation matrix} \\ \begin{array}{cc} \textit{Present value} & \textit{Proposed change} \end{array} \\ \begin{pmatrix} 45 & -9 \\ 41 & -7 \\ 65 & -10 \end{pmatrix} \end{array}$$

This process of row into column multiplication is quite general, and has certain results with which the reader is probably familiar, viz.

(*i*) Apart from any informational sense or nonsense it is quite impossible to multiply two matrices **A**, **B** set down in this order unless they are *conformable* for this multiplication; i.e. the row length of **A** equals the column length of **B**. For example, it is impossible to multiply:

$$\begin{pmatrix} 15 & -3 \\ 13 & -2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 1 & 2 \\ 0 & 5 \end{pmatrix}.$$

(*ii*) Even if multiplication both ways is possible—what is the condition for this?—the product matrices may not be equal.

For example:

$$\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 14 \\ -3 & 12 \end{pmatrix}$$

$$\text{but } \begin{pmatrix} 3 & 5 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 18 \\ -2 & 11 \end{pmatrix}.$$

The student should make up an example and verify that usually $\mathbf{AB} \neq \mathbf{BA}$.

9.2 An example of the use of matrices, which may be regarded at first as a mere convenience, is to write a pair of simultaneous equations in the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \qquad \text{(equation E).}$$

We shall take the opportunity to state the rule for multiplication of matrices quite formally at this point, viz. the inner product of the 1st row (of **A**) with the 1st column (of **B**) is the term in the 1st row and 1st column of the product (**AB**); and in general, the inner product of the ith row of **A** with the jth column of **B**, is the element in the ith row and jth column of the product **AB**.

Thus: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$, and from equality of vectors we have $\begin{matrix} ax+by = p \\ cx+dy = q \end{matrix}$.

Returning to the equation E, we see that the matrix form emphasises a new aspect of the situation. The 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the pattern of left-hand-side coefficients. If we have a pair $\begin{pmatrix} x \\ y \end{pmatrix}$ of values of the variables, supposed to be a solution of the equations, then the pattern of coefficients specifies the process (substitution) by which we operate on the solutions to get, we hope, the pair of values $\begin{pmatrix} p \\ q \end{pmatrix}$. If we write **M** for the matrix, and **v**, **k** for the vectors $\begin{pmatrix} x \\ y \end{pmatrix}$ and $\begin{pmatrix} p \\ q \end{pmatrix}$ respectively, then $\mathbf{Mv} = \mathbf{k}$.

This equation highlights the action of **M** in multiplying **v** as a piece of operating machinery which converts **v** into another vector **k**—just as, at a simpler level, we may regard the '5' in the expression 5**v** as an operating machinery which converts **v** into another vector which is in the same direction but five times the size.

EXERCISE 9

(1) Find a simple matrix equal to the following products:

(*a*) $\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 6 \\ 0 & 2 \end{pmatrix}$ (*b*) $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$

(*c*) $\begin{pmatrix} 1 & 6 \\ 2 & 13 \end{pmatrix} \begin{pmatrix} 13 & -6 \\ -2 & 1 \end{pmatrix}$ (*d*) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$

(*e*) $(-1 \ \ -2 \ \ 1) \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$ (*f*) $\begin{pmatrix} 1 & 6 \\ 2 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$

(*g*) $\begin{pmatrix} 13 & -6 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 6 \\ 2 & 13 \end{pmatrix}$ (*h*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 2 \\ 2 & 3 & -5 \\ 6 & 2 & 1 \end{pmatrix}$

(*i*) $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$.

(2) If $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, evaluate **AB** and **BC**. Then calculate (**AB**)**C** and **A**(**BC**).

(3) If $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 6 \\ 1 & 1 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ evaluate **AB**, **AC**, **B**+**C**, **A**(**B**+**C**). (We add matrices by adding corresponding elements.) Is **AB**+**AC** = **A**(**B**+**C**)?

(4) If $\mathbf{A} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 6 \\ 1 & -2 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \\ -1 & 6 \end{pmatrix}$, say which of the following products exist, and evaluate those.

(*i*) **AB**	(*iii*) **CA**	(*v*) **BC**
(*ii*) **BA**	(*iv*) **AC**	(*vi*) **CB**

(5) A matrix is described as of shape $m \times n$ (m by n) if it has 'm elements down and n across'; e.g. in question 1(*b*) the matrices multiplied are 2×2 and 2×1. Specify the shapes in 1(*d*) (*e*) (*f*). Give a rule (*i*) for matrices ($m \times n$) ($p \times q$) to be conformable for multiplication in that order, and (*ii*) for the shape of the result.

(6) If $\mathbf{A} = \begin{pmatrix} 2 & 6 \\ 1 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, evaluate **AB** and **BA**. Do these matrices commute?

(7) If $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ -1 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 5 & 3 \\ 3 & -1 \end{pmatrix}$ and $\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, compute **AX** and **BX**. Notice $\mathbf{AX} = \mathbf{BX} \not\Rightarrow \mathbf{A} = \mathbf{B}$. Can you find another algebraic structure for which such a cancellation rule does not hold? (*Hint*: Try set algebra and modular arithmetics, e.g. modulo 4 and multiplication.)

(8) If $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 4 & 1 \\ 3 & -1 & 2 \\ -5 & -3 & -3 \end{pmatrix}$ compute **AB**.

Can you find another algebraic system in which $\mathbf{AB} = \mathbf{0} \not\Rightarrow \mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$ or both? (**0** stands for a matrix in which every term is zero.)

(9) Fill in the gaps in the following matrix equations.

$$(i) \quad \begin{pmatrix} 5 & 10 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} . & 0 \\ . & . \end{pmatrix} \begin{pmatrix} 1 & . \\ 0 & 1 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} 3 & -3 & 6 \\ -1 & 5 & 14 \\ -3 & 1 & 12 \end{pmatrix} = \begin{pmatrix} . & 0 & 0 \\ . & . & 0 \\ . & . & . \end{pmatrix} \begin{pmatrix} 1 & . & . \\ 0 & 1 & . \\ 0 & 0 & 1 \end{pmatrix}$$

Factorisation in this way is the basis of Choleski's method—a very efficient one—for simultaneous equations. Matrices of the type on the right-hand side are called 'triangular'.

(10) There are contexts in which a row 3-vector and a column 3-vector both appear, e.g. $(p\ q\ r)$ and $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Form the product $(p\ q\ r) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ by the matrix rule.

Can the result be called a 'matrix'? Have you already met a process similar to this for two vectors of the same order?

(11) Premultiply the matrix $\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ by each of the following matrices.

Describe in words in each case what the multiplier does to the rows of M.

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(12) Find the product of these matrices.

$$(x\ y\ 1) \begin{pmatrix} 0 & 0 & -2a \\ 0 & 1 & 0 \\ -2a & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

What would be obtained by putting this product equal to zero?

chapter 10

MAPPINGS

10.1 At the end of the last chapter we made a remark to the effect that when a matrix multiplies a vector, the result is to transform this vector into another vector. A transfer operation such as this is called a *mapping*, and we make a diversion to study in some detail the general idea of a mapping.

The term 'mapping' is in regular use in the language of sets: it is a relation between two sets (they may be the same sets) subject to certain rules. If we picture the sets as being collections of dots in the following diagrams, they exhibit what is and what is not a mapping.

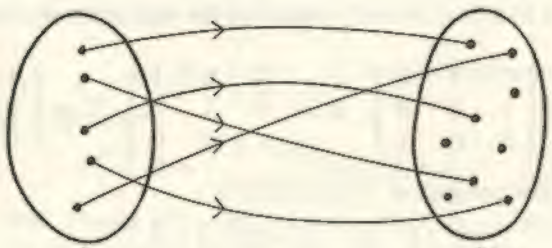


FIG. 36

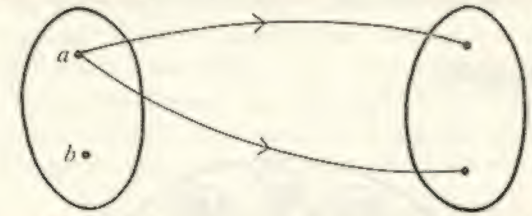
A mapping of S into S′†

N.B. (*i*) Every member of S has just one image in S′, although some images may coincide.

(*ii*) It is possible for some members of S′ not to be images.

† The fact that the elements of S are represented by dots is inclined to give the impression that there is necessarily a finite number of elements in S. This is not so, as exercise 6a shows.

The following is not a mapping because it violates condition (*i*) in two ways.



a has two images, b has none

FIG. 37

10.2 Mappings themselves may be divided into two classes, illustrated by the two diagrams:

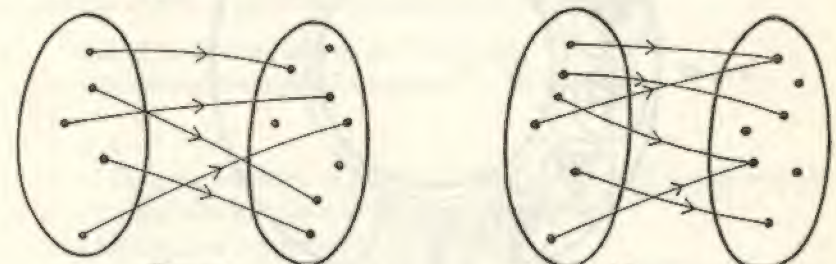


FIG. 38

The characteristic which distinguishes these two mappings is that, in the first, each element of S is mapped on to a different element of S′, whereas no such restriction exists for the second. For obvious reasons the first is called a one-one mapping, and the second a many-one mapping.

However, for neither of these mappings is it possible to reverse the direction of the arrows to get a mapping of S′ into S—condition (*i*) would not be satisfied. It is clear then that for a mapping to be reversible, there must be no spare elements in S′; in fact the diagram must be as follows:

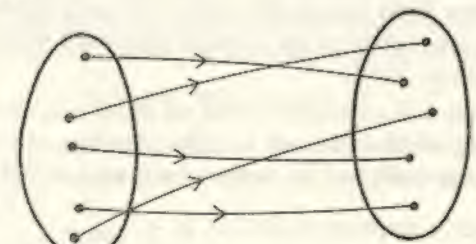


FIG. 39

An example of a reversible one-one mapping

For any mapping, the set of object elements is called the *domain*, and the set which includes the image elements is called the *range* (or *co-domain*).

10.3 *Examples of mappings*

(*a*)

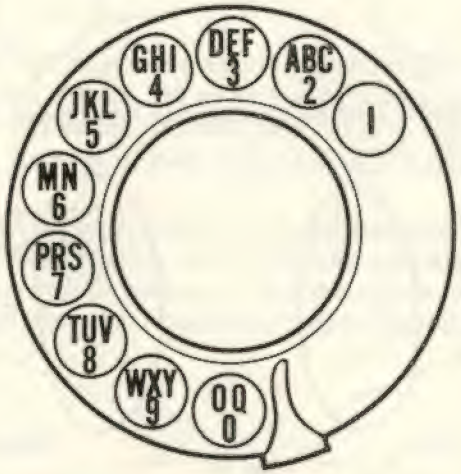


FIG. 40

The dial of the current London telephone is shown in the diagram. This can be regarded as a mapping of the letters—if we exclude Z—into the numbers. Given one of the letters, say N, there is a number corresponding, 6. It is *not* a mapping from the numbers to the letters. Give two reasons why it is not.

(*b*) *Fingerprints*

Scotland Yard has a file of fingerprints, and to each set of fingerprints there corresponds one (and only one) person. This is a one-one mapping, with the set of fingerprints as the domain, and the set of people as the range. Note that in this case, not every element of the range (person) has an element of the domain associated with it (has his fingerprints in Scotland Yard).

(*c*) $x \to 2x+3$ where x is a member of the set of all real numbers.

This is a mapping which is defined for all real values of x. It is also reversible, for we can work out the value of x if we are told the value of $2x+3$. The reverse mapping is, in fact, $x \to \dfrac{x-3}{2}$.

EXERCISE **10a**

In each of the following questions, decide whether or not we have defined a mapping. If it is a mapping, say whether it is many-one or one-one, and also if it is reversible; if it is not a mapping, say so. If, by suitably restricting the domain, you can make a non-mapping into a mapping, say so. (Assume that x may be any member of the set of real numbers. We use the brace notation when specifying sets.)

Example: $x \to \dfrac{1}{x}$ is not a mapping since there is no image of the real number zero. We can make it a mapping by restricting the domain so that for $x \neq 0$, $x \to \dfrac{1}{x}$ *is* a one-one, reversible mapping.

Discuss, as detailed above, the following:

(1) {Citizens of the United Kingdom} → {Surnames}.

(2) $x \to x^{\frac{1}{2}}$

(3) $x \to x^2$

(4) {Telephone numbers in the United Kingdom} → {Citizens responsible for the bills for the individual telephones}.

(5) {Articles in a supermarket} → {Prices}.

(6) {Male citizens of United Kingdom over 21} → {The banks which they use}.

(7) A complex number $Z \to$ real number $|Z|$.

(8) 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to$ values of $(ad-bc)$.

(9) $x \to$ angle which has a sine equal to x.

(10) $x \to \log x$.

10.4 *Composition of mappings*

Consider the following two mappings performed upon the set of real numbers. Letting x be any member of this set, we suppose that the mapping **S** changes the sign of x and we write this: $\mathbf{S}x = -x$.

R is the operation 'square x', so that we write $\mathbf{R}x = x^2$.

If we wished to follow the mapping **S** by **R** we would write:

$$\mathbf{R}(\mathbf{S}x) = \mathbf{R}(-x) = (-x)^2 = x^2$$

We use the expression $\mathbf{R}(\mathbf{S}x)$—the result of following mapping **S** by mapping **R**—to *define* the product of the mappings **RS**, so that $(\mathbf{RS})x$ means $\mathbf{R}(\mathbf{S}x)$.

We see that $(\mathbf{RS})x = x^2$.

Note (*a*) by virtue of the definition, there is no ambiguity in writing **RS**x without brackets;

(*b*) that **RS** means that we perform the mapping **S** and then follow it by **R**.†

For the same example,

$$(\mathbf{SR})x = \mathbf{S}(\mathbf{R}x) = \mathbf{S}x^2 = -x^2,$$

so we see that the mapping **SR** is not the same as the mapping **RS**. Sometimes we may find that mappings commute, but these are special cases.

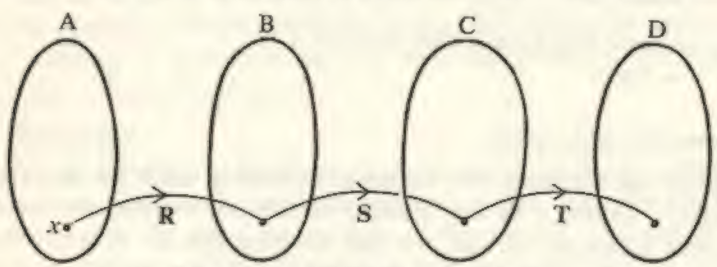
For these mappings we see that $\mathbf{R}x = x^2$ and $\mathbf{RS}x = x^2$, making it tempting to say $\mathbf{R} = \mathbf{RS}$. We allow this if and only if the effect of **R** and **RS** is the same for every element. This is clearly so, and we may say $\mathbf{R} = \mathbf{RS}$.

Finally we note that $\mathbf{S}.\mathbf{S}(x) = \mathbf{S}(\mathbf{S}x) = \mathbf{S}(-x) = x$.

The mapping **S.S** is often written $\mathbf{S}^2$ to fit usual multiplication practice, and we notice that the effect of $\mathbf{S}^2$ is to leave x unchanged—this is called the identity mapping with the letter **I** reserved for it. In this case, $\mathbf{S}^2 = \mathbf{I}$.

10.5 *Associative rule for mappings*

If we think of three mappings performed successively on a set of elements we shall have a situation as shown in the following diagram.



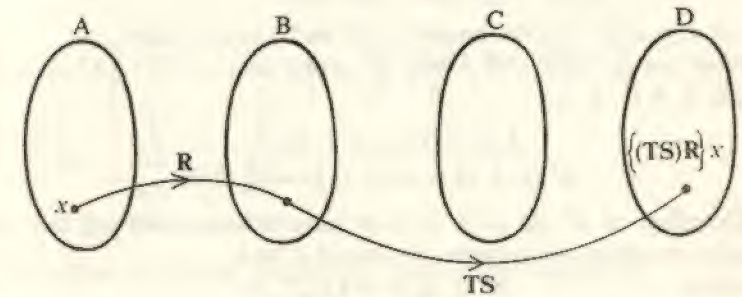
The Associative Rule for Mappings

FIG. 41

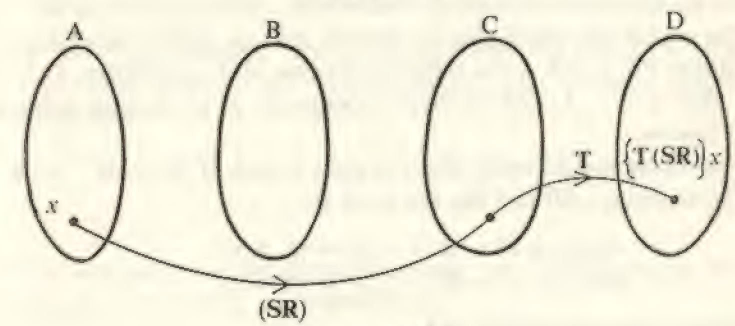
The set A is mapped into B by **R**, into C by **S** and then into D by **T**. We show one element from A mapped successively into B, C and D and we call that element x. The diagram shows the final result: $\mathbf{T}\{\mathbf{S}(\mathbf{R}x)\}$.

† We leave it to the reader to convince himself that **RS** is itself a mapping.

By definition of product of mappings we can write this in two ways—$(\mathbf{TS})(\mathbf{R}x)$ or $\mathbf{T}(\mathbf{SR})x$. The first of these alternatives we would call the mapping **(TS)R** and the second **T(SR)**. These are apparently equivalent mappings. Each can be represented diagrammatically, as follows:



Mapping **(TS)R**



Mapping **T(SR)**

The Associative Rule for Mappings

FIG. 42

It is as if we wish to travel from Aberdeen to Dover via Bradford and Cambridge, and there are two coach firms offering package deals. The first drives to Bradford, and then by-passes Cambridge making straight for Dover: the second drives to Cambridge, by-passing Bradford, and then proceeds to Dover. But it does not matter whether we take package deal **(TS)R** or **T(SR)**: we still get to Dover!

In view of the fact that there is no ambiguity in the mapping **TSR**, we shall no longer put in the brackets.

10.6 *A worked example of composition of mappings*

Consider the set of arrangements of the numbers 1, 2 and 3, and let **A** and **B** be mappings of this set on to itself, defined as follows:

$$\mathbf{A}(x, y, z) = (y, x, z)$$
$$\mathbf{B}(x, y, z) = (x, z, y),$$

where x, y, and z are the numbers 1, 2 and 3 in any order.

What are the effects of **A** and $\mathbf{A}^2$ acting on (1, 2, 3)? [$\mathbf{A}^2$ (1, 2, 3) means $\mathbf{A}.\mathbf{A}$ (1, 2, 3).]

$$\mathbf{A}(1, 2, 3) = (2, 1, 3)$$
$$\mathbf{A}^2(1, 2, 3) = \mathbf{A}(2, 1, 3) = (1, 2, 3)$$

The effect of $\mathbf{A}^2$ on (1, 2, 3) is to leave it alone—we call this the identity mapping and as usual we denote it by **I**.

Hence, $\mathbf{A}^2(1, 2, 3) = \mathbf{I}(1, 2, 3)$,
and, in general, $\mathbf{A}^2(x, y, z) = \mathbf{I}(x, y, z)$.

We discussed this situation in section 10.4—we may now write $\mathbf{A}^2 = \mathbf{I}$.

The reader will recall that an element P of an algebra satisfying the condition $PA = AP = I$ is called the inverse of A and written A^{-1}. In this case $\mathbf{A}^{-1} = \mathbf{A}$, and we have an example of an element being its own inverse.

The reader should verify that the same is true of **B**, i.e. $\mathbf{B}^{-1} = \mathbf{B}$.

The mappings **AB** and **BA** are given by:

$$\mathbf{AB}(x, y, z) = \mathbf{A}(x, z, y) = (z, x, y)$$
$$\mathbf{BA}(x, y, z) = \mathbf{B}(y, x, z) = (y, z, x)$$

Hence, once again $\mathbf{AB} \neq \mathbf{BA}$.

Let us try to build up from **AB**, and see what further mappings we can manufacture.

First we consider **AAB** (x, y, z). Because the associative law holds for mappings we can say:

$$\mathbf{AAB}(x, y, z) = \mathbf{B}(x, y, z) = (x, y, z).$$

This is not a new mapping, so we try **BAB** (x, y, z).

Now $\mathbf{BAB}(x, y, z) = \mathbf{B}(z, x, y) = (z, y, x)$.

This is a new mapping, and we add it to the list, now consisting of **I**, **A**, **B**, **AB**, **BA**, and **BAB**.

Let us try building up from **BA**. $\mathbf{BBA} = \mathbf{IA} = \mathbf{A}$ and we get no further. Also, $\mathbf{ABA}(x, y, z) = \mathbf{A}(y, z, x) = (z, y, x)$, and we see that $\mathbf{ABA} = \mathbf{BAB}$.

In fact, if we proceed further we find that there are no more mappings that we can find—they have the property of closure under the operation 'followed by'.

The reader is invited to copy and complete the table below:

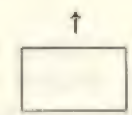
(*First operation*)

(*Second operation*)		I	A	B	AB	BA	BAB
	I	I	A	B	AB	BA	BAB
	A	A	I	AB	B		
	B	B					
	AB	AB					
	BA	BA					
	BAB	BAB					

Note that, as always, we *write* first the *second* operation.

EXERCISE 10b

(1) A squad, consisting of 15 soldiers, falls in (at the start of a parade) in three ranks facing in a certain direction. This is shown diagrammatically thus:



If the sergeant gives the order 'left wheel' (or 'left form') the whole squad turns left en bloc, to become thus:

This operation will be symbolised by **W**. If, however, he gives the order 'left turn', **L**, each man individually turns left, to give this result:

Explain why an officer coming on parade cannot distinguish between the results of **LL** and **WW**, nor between $\mathbf{L}^2\mathbf{W}^2$ and **I** (= no order has been given). Give other combinations equivalent to **I**, first involving **L** and **W** only and secondly involving other orders of your own choosing.

(*Note*: The whole range of formations involving right angles can be derived from **L** and **W**.)

(2) Let **V**, **D**, **S** be mappings of the set of real numbers such that $\mathbf{V}(x) = 1-x$, $\mathbf{D}(x) = 2x$ and $\mathbf{S}(x) = x^2$. Find the mappings **DV**, **VS**, and **SV** and give the particular values of x for which **VS** = **SV**. Find $\mathbf{D}(\mathbf{VS})(x)$ and $(\mathbf{DV})\mathbf{S}(x)$ and hence show that these mappings are associative. What is the inverse of **V**?

(3) We define mappings on the set of real numbers excluding zero as follows:

$$\mathbf{f} : x \to -x$$
$$\mathbf{g} : x \to 1/x$$
$$\mathbf{i} : x \to x.†$$

What are the mappings $\mathbf{f}^2$, $\mathbf{g}^2$, **fg** and **gf**? Use your expressions for these mappings to simplify **fgf** and **gfg**. Copy and complete the table:

		i	f	g	fg	(*First operation*)
(*Second operation*)	i	i	f	g	fg	
	f	f				
	g	g				
	fg	fg				

(4) For x real and positive we define $\mathbf{l}:x \to \log x$, $\mathbf{r}:x \to x^{\frac{1}{2}}$ and $\mathbf{h}:x \to \frac{1}{2}x$. Show that **hl** = **lr**. What is the usual algebraic way of writing $\mathbf{l}^{-1}$? Rewrite $\mathbf{r} = \mathbf{l}^{-1}\mathbf{hl}$ using algebraic symbols and notation.

(5) Let $\mathbf{i}:x \to x$, $\mathbf{a}:x \to 1-x$, $\mathbf{b}:x \to \dfrac{1}{x}$, be mappings defined for x belonging to set of real numbers ($x \neq 0$ or 1).

Try to build up a table of form:

		i	a	b	etc
Do this operation second.	i	i	a	b	
	a	a			
	b	b	ba		
	etc				

printing in as many different mappings as you need. You should find it necessary to have 6 altogether. (*Hint*: Put **c** = **ba**.)

† This way of writing mappings is common. We read $f{:}x \to -x$ as the mapping f such that x is mapped into $-x$.

(6) If $y(x) = \dfrac{1-x}{1+x}$, what is $y^2(x)$?

(7) Bell ringers ringing on 4 bells start with the bells in order 1 2 3 4 —called 'rounds'—and aim to ring all the possible permutations (changes) in order, ending up with rounds, subject to certain rules. These are:

(*a*) No bell may ring in the same position more than twice in succession.

(*b*) Each bell may move at most one place at a time from one change to another.

If a, b, c, d are the numbers 1, 2, 3, 4 in any order, bell ringers define the following mappings:

From a b c d

to b a d c will be written **x**.

From a b c d

to a c b d will be written **y**.

Show that $\mathbf{x}^2 = \mathbf{y}^2 = \mathbf{i}$ where **i** is the identity.

Writing **z** for the mapping **yx**, show that $\mathbf{z}^n = \mathbf{i}$ for a certain value of n and its multiples. Find this value, and show $\mathbf{x}\mathbf{z}^{n-1} = \mathbf{y}$.

We define another operation **w** which is a b c d

a b d c.

Solve the equation $(\mathbf{wy})^m = \mathbf{i}$. Show how **x**, **y** and **w** can be used to ring all 24 changes.

(8) Two transformations† of the set of complex numbers are defined by $\mathbf{T}z = -\dfrac{1}{z}$, $\mathbf{S}z = z+1$. Show that $\mathbf{TST} = \mathbf{S}^{-1}\mathbf{TS}^{-1}$, $\mathbf{STSTST} = \mathbf{TSTSTS} = \mathbf{I}$, where $\mathbf{I}z = z$ is the identical transformation and $\mathbf{S}^{-1}z = z-1$ is the inverse of the transformation **S**.

(*Cambridge Scholarship*)

† A transformation is the same as a mapping.

(9) The following transformations are defined on the set of complex numbers:

(*i*) $R(z) = jz-j+1$; (*ii*) $S(z) = z/(z-1)$.

Show that $\mathbf{R}^4 = \mathbf{S}^2 = \mathbf{I}$, where $\mathbf{I}$ is the identity transformation $\mathbf{I}(z) = z$, and show also that $\mathbf{RS} = \mathbf{SR}^3$.

Deduce that there are eight, and only eight, distinct transformations which can be obtained by arbitrary (finite) combinations of the transformations **R** and **S**. (*Cambridge Scholarship*)

chapter 11

MAPPINGS OF A VECTOR SPACE

11.1 The reader will recall that, in a vector space, all vectors can be expressed in terms of a basis by taking the suitable multiples and adding them. If **u** and **v** are the vectors of any basis, then any vector of the space may be written uniquely in the form $p\mathbf{u}+q\mathbf{v}$.

To study mappings of a vector space, we shall consider for simplicity mapping a plane, which may be visualised as a piece of graph-paper extended indefinitely. In view of the grid it is natural to use the unit vectors **i** and **j**, parallel to these lines, as a basis. We shall also, for simplicity, picture representative vectors drawn from a chosen origin O. Then for a typical point P with co-ordinates x and y we have a vector $\begin{pmatrix} x \\ y \end{pmatrix}$, expressible also as $\mathbf{v} = x\mathbf{i}+y\mathbf{j}$, which is the position vector of P with respect to O. Quite often we shall prefer to talk in terms of the points rather than the vectors since it will then be easier to visualise what is happening when we perform a mapping.

In this way we can imagine a given mapping as a process of carrying points from one location to another. It is immaterial whether we picture the image point on a new piece of graph paper, or superimposed on the same piece. For clarity, in the early stages we use the former.

11.2 We can now talk about a mapping of a plane S into another plane S' in this way: we suppose that the vectors of a basis of S, viz. **u** and **v**, are mapped into vectors **U** and **V** which are in S'. A particularly simple kind of mapping is one in which $p\mathbf{u}+q\mathbf{v} \longrightarrow p\mathbf{U}+q\mathbf{V}$, i.e. a linear combination of **u** and **v** is mapped into the same linear combination of **U** and **V**. Once this is known every vector of the vector space may be very simply mapped into its image. Such a mapping is called a *linear mapping* or a *linear transformation*. (*Note*: **U** and **V** need not themselves form a basis for a plane-system, for there is no safeguard against them being linearly dependent. There is also no need for **U** and **V** to be the same 'size' vectors as **u** and **v**, as example (1) below will show.)

Usually, of course, we shall take **i** and **j** as our basis for a space of two dimensions, and **i**, **j** and **k** as our basis for a space of three dimensions.

11.3 Let us consider some mappings.

(*a*) *Transformation corresponding to a geographical map*

We consider an ordinary map, scale 1 inch to 1 mile, which is correctly aligned with the countryside so that no rotation takes place. Also we select an origin on the ground with a corresponding origin on the map, and anchor our vectors there so that the vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ for the ground represents a point, and $\begin{pmatrix} x' \\ y' \end{pmatrix}$ for the map also represents a point. If z is measured vertically upwards, the required mapping is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longrightarrow \begin{pmatrix} x' \\ y' \end{pmatrix}$, where $x' = \dfrac{x}{N}$, $y' = \dfrac{y}{N}$ and $N = 63{,}360$.

Notice that this is a mapping of 3-vectors into 2-vectors, with the z co-ordinate being suppressed completely.

The basis vectors of the countryside **i**, **j** and **k** are mapped into $\dfrac{1}{N}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\dfrac{1}{N}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Any vector of the 3-space $x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ is mapped into the corresponding multiples of the images of the basis vectors, i.e. $x\mathbf{i}+y\mathbf{j}+z\mathbf{k} \longrightarrow \dfrac{x}{N}\begin{pmatrix} 1 \\ 0 \end{pmatrix}+\dfrac{y}{N}\begin{pmatrix} 0 \\ 1 \end{pmatrix}+z\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

It can then be seen that this is a linear mapping.

(*b*) *The one-way stretch*

This is the mapping in which a vector space is stretched in one direction only: we consider the simple case in which stretch is in the x-direction by a factor of 2, i.e. $\begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow \begin{pmatrix} 2x \\ y \end{pmatrix}$.

The vector $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is transformed to $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ which we call **a** in the transformed space; and $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ has the transform $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{b}$, using a

new name to emphasise that it also is an image. The definition of a linear mapping is that for all values of p, q the vector $p\mathbf{i}+q\mathbf{j} \rightarrow p\mathbf{a}+q\mathbf{b}$, and we now test whether this is true for this particular case. Now $p\mathbf{i}+q\mathbf{j} = p\begin{pmatrix} 1 \\ 0 \end{pmatrix}+q\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$ and this transforms to $\begin{pmatrix} 2p \\ q \end{pmatrix}$ directly.

Testing the formula, we have $p\mathbf{a}+q\mathbf{b} = p\begin{pmatrix} 2 \\ 0 \end{pmatrix}+q\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2p \\ q \end{pmatrix}$, agreeing with the previous result.

Pictorially, the image builds up on the transforms **a**, **b** just as the original builds on **i** and **j**:

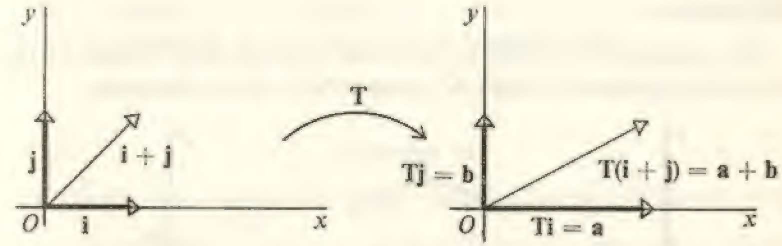


FIG. 43

(*c*) *Change of origin*

We consider the points of the plane as if they were on a piece of plain paper, with the axes on a piece of tracing paper above it. The bottom piece of paper is slid from under the axes through 2 units in the x-direction while the origin and the axes remain fixed in space. Each vector which joined the origin to a point $P(x, y)$ now has its x-component increased by 2 units, i.e. $\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x+2 \\ y \end{pmatrix}$.

In particular, $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \mathbf{a}$ and $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{b}$.

Considering any vector expressed in terms of the basis, we have $p\mathbf{i}+q\mathbf{j} = \begin{pmatrix} p \\ q \end{pmatrix} \rightarrow \begin{pmatrix} p+2 \\ q \end{pmatrix}$, while the corresponding expression in terms of **a** and **b** is $p\mathbf{a}+q\mathbf{b} = \begin{pmatrix} 3p \\ q \end{pmatrix}$.

Since these are not equal, the transformation is not linear. (*Note*: If we happen to choose $p = 1$, the discrepancy will not show.)

Fig. 44 illustrates the case $p = -1$, $q = 1$, i.e. it transforms the vector $-\mathbf{i}+\mathbf{j}$.

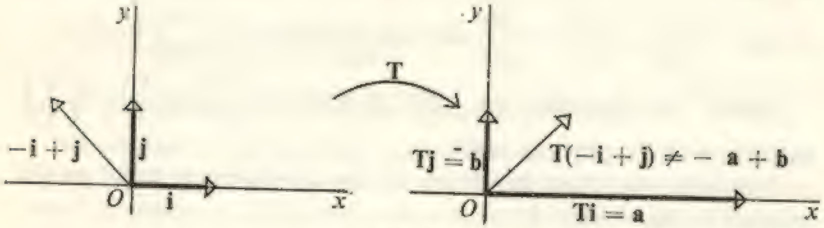


FIG. 44

(*d*) *Rotation*

The mapping **R** we define as being the mapping which rotates all the vectors in the plane through 90° in an anti-clockwise direction.

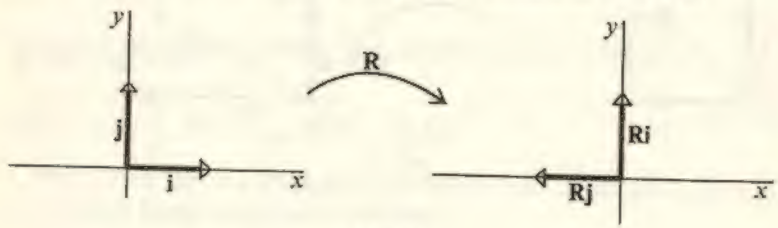


FIG. 45

We see that $\mathbf{i} \to \mathbf{j}$ and $\mathbf{j} \to -\mathbf{i}$. Also any point $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} -y \\ x \end{pmatrix}$ (see diagram 46). The basis vectors are transformed into $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ respectively; and $\begin{pmatrix} x \\ y \end{pmatrix} = x\mathbf{i}+y\mathbf{j} \to x\begin{pmatrix} 0 \\ 1 \end{pmatrix}+y\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$.

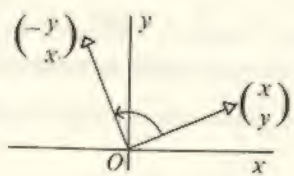


FIG. 46

Hence rotation about the origin *is* a linear mapping.

11.4 *Worked example of a mapping*

We discuss whether the mapping of points of a plane into a line, given by $\begin{pmatrix} x \\ y \end{pmatrix} \to (xy+y)$, is a linear mapping or not. This would not be an easy mapping to represent pictorially—we discuss it algebraically.

We take as basis for the plane as usual **i** and **j**, and we see that $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \to (0)$ and $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \to (1)$.

If this were a linear mapping, $\begin{pmatrix} x \\ y \end{pmatrix}$, which is $x\mathbf{i}+y\mathbf{j} \to x$ (image of **i**) $+y$ (image of **j**).

But x (image of **i**)$+y$ (image of **j**) $= x(0)+y(1) = (y)$, and because this is not equal to $(xy+y)$ this is *not* a linear mapping.

EXERCISE 11

(1) The plane is mapped on to the x-axis by joining each point of the plane to the point $\begin{pmatrix} 0 \\ a \end{pmatrix}$ and producing this line (if necessary) to cut the x-axis. Find the point into which the point $\begin{pmatrix} x \\ y \end{pmatrix}$ maps, and also the images of **i** and **j**, and decide whether $x\mathbf{i}+y\mathbf{j}$ maps into x (image of **i**)$+$ y (image of **j**), i.e. whether the mapping is linear.

(2) Test the following mappings for linearity:

(*a*) $x \to kx$ (*b*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ 2y \end{pmatrix}$ (*c*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 1/x \\ 1/y \end{pmatrix}$

(*d*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{pmatrix}$ (*e*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x^2+y^2 \\ 0 \end{pmatrix}$ (*f*) $\begin{pmatrix} x \\ y \end{pmatrix} \to (xy)$

(*g*) $\begin{pmatrix} x \\ y \end{pmatrix} \to (x+y)$ (*h*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ y \\ x+y \end{pmatrix}$ (*i*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$

(3) Let the components of a vector be the coefficients of a quadratic polynomial, so that ax^2+bx+c corresponds to $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Let **D** be the mapping which differentiates ax^2+bx+c with respect to x. Is this mapping linear?

(4) The following transformation is sometimes used to prove that there are exactly as many points on the x-axis between 0 and 1 as there are in the unit square with $\mathbf{i}$ and $\mathbf{j}$ as adjacent sides. Take any point on the x-axis and find the number corresponding to it, say, 0·3712312. Take alternate digits, make up new numbers 0·3132 and 0·721, and take these as the x and y co-ordinates of a point in the unit square. Is the mapping linear?

(5) Taking x, y and z axis in the east, north and vertically upwards directions, and supposing that the sun is in the x-z plane at an angle of α° with the positive x-direction, what is the equation of the mapping of each point in space on to its shadow on the plane? Is this mapping linear?

(6) We now prove that a linear mapping of a vector 2-space into a vector 2-space must be of a certain form, which the reader will probably have anticipated after working question 2.

We suppose that $\mathbf{i} \to \begin{pmatrix} a \\ c \end{pmatrix}$ and $\mathbf{j} \to \begin{pmatrix} b \\ d \end{pmatrix}$. If the mapping is linear, what is the image of $x\mathbf{i}+x\mathbf{j}$? What is this, written as a column vector?

If we say $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix}$, what are the equations for x' and y'? (*Note*: We have here proved that if a mapping is linear it is of the form $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$, i.e. that $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$. In question 2(*i*) we proved the converse.)

chapter 12

MATRIX MAPPINGS OF A VECTOR SPACE

12.1 We now consider the mapping of a vector space which is carried out by a matrix. Take any 2×2 matrix $\mathbf{M}$, and transform $\mathbf{v}$ by (pre-)multiplying by $\mathbf{M}$ to obtain an image vector $\mathbf{v}'$.

Then, $\mathbf{Mv} = \mathbf{v}'$ or in full, $\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$ where $\mathbf{M} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$.

For the reader who has studied transformation geometry (sometimes called 'motion geometry') the following special transformations will be familiar, but the emphasis in geometry is of course on the transformation of all the *points* to positions on another plane—this plane usually being taken as coincident with the first.

Special cases of 2×2 transformation matrices

(*i*) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$

The matrix is called the identity matrix, or the unit matrix of order 2, symbolised by $\mathbf{I}$. It leaves all 2-vectors unchanged, including the basis vectors as special cases.

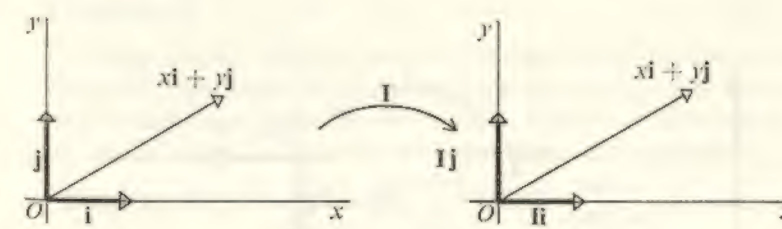


FIG. 47

(*Note*: From now on, we shall carry out our matrix transformations first on $\mathbf{i}, \mathbf{j}$; we consider the general case $x\mathbf{i}+y\mathbf{j}$ or $\begin{pmatrix} x \\ y \end{pmatrix}$ later.)

(*ii*) $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} k \\ 0 \end{pmatrix} = k\mathbf{i}$; $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{j}$, i.e. if the matrix is T then $\mathbf{Ti} = k\mathbf{i}$, $\mathbf{Tj} = \mathbf{j}$.

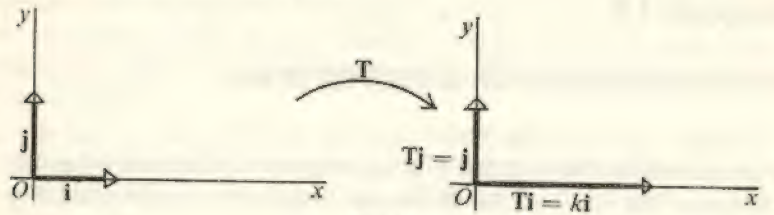


FIG. 48

(*iii*) We could now consider $\begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}$. This is left to the reader with note that it includes the important special case $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$.

(*iv*) Consider a 2×2 matrix in which the first column is $\begin{pmatrix} p \\ q \end{pmatrix}$, the other elements being unknown. (They must of course have definite values, if the mapping is to be valid and determined.) Then, $\mathbf{Ti} = \begin{pmatrix} p & . \\ q & . \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$. $\mathbf{Tj}$ cannot be stated.

However, *the first column of the matrix states the transform of* **i**. The reader should verify that the second column shows the transform of **j**.

We can now construct matrices which transform **i** and **j** into specified positions.

(*v*) To reflect **i** and **j** in the x-axis:

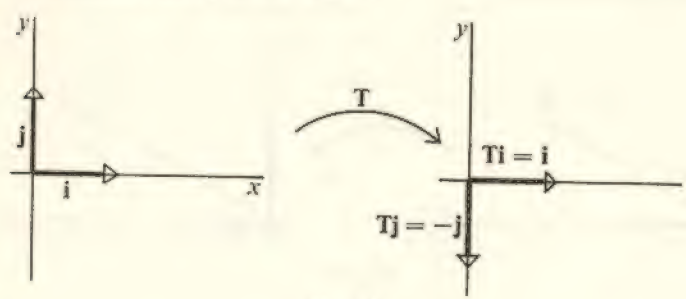


FIG. 49

We have $\mathbf{Ti} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{Tj} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and the matrix is then $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(*vi*) To reflect **i** and **j** in the line $x+y = 0$:

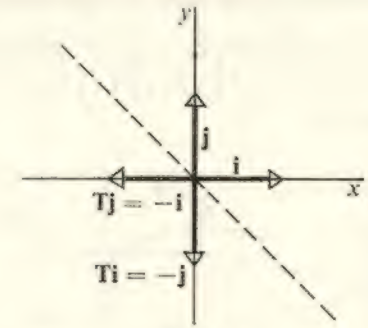


FIG. 50

$\mathbf{Ti} = -\mathbf{j} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$; $\mathbf{Tj} = -\mathbf{i} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. The required matrix is $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$.

(*vii*) To reflect **i** and **j** in the line $y = x \tan \alpha$:

Ti is of length unity at an angle 2α to Ox.

Hence $\mathbf{Ti} = \begin{pmatrix} \cos 2\alpha \\ \sin 2\alpha \end{pmatrix}$ and $\mathbf{Tj} = \begin{pmatrix} -\sin 2\alpha \\ \cos 2\alpha \end{pmatrix}$.

The matrix is therefore $\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$. Verify that $\alpha = 45°$ gives one of (*v*) (*vi*) and that $\alpha = -45°$ gives the other.

12.2 *A useful check*

It is clear that in reflection, as in some other transformations which follow, the right angle which the basis vectors **i** and **j** make with each other is unchanged by the transformation. If this is so, then the resulting vectors should be tested for orthogonality.

For example, in (*vii*): $\mathbf{i} \to \begin{pmatrix} \cos 2\alpha \\ \sin 2\alpha \end{pmatrix}$; $\mathbf{j} = \begin{pmatrix} \sin 2\alpha \\ -\cos 2\alpha \end{pmatrix}$, and the inner product of these two is easily seen to be zero—they are still orthogonal.

The transformation of only the vectors **i** and **j** from among the infinite set of vectors in the space would be of little interest; however, it is immediately apparent that we can make a number of deductions when we know the transforms of these particular vectors.

EXERCISE **12a**

(1) Draw a diagram showing the effect of the matrix $\mathbf{M} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ on the points $\begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}2\\0\end{pmatrix}\begin{pmatrix}3\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix}\begin{pmatrix}0\\2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix}$, i.e. a diagram of the form:

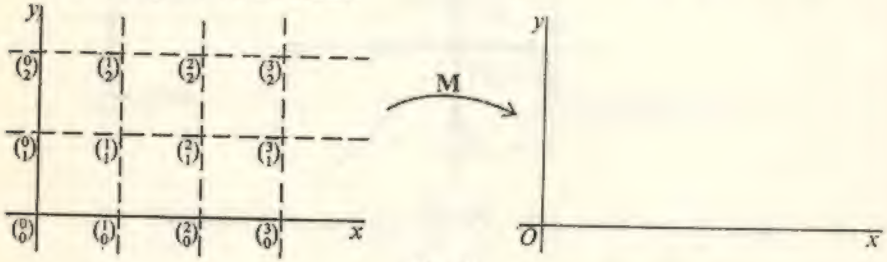


FIG. 51

(2) Illustrate the effect of the matrix $\mathbf{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ upon the square with **i** and **j** as adjacent sides. Write down in words what operation is performed on the unit square. Evaluate the matrix $\mathbf{M}^2$ and illustrate in the same way. Evaluate $\mathbf{M}^4$ and explain the result from your diagrams.

(3) Draw diagrams to illustrate the operating effects of $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$, and describe these mappings separately. Without multiplying the matrices, but using your word descriptions, do the mapping represented by **A** followed by the mapping for **B**, deciding what your final result is. Verify by multiplying directly to form the matrix **BA** that what you decided was right.

(4) Investigate the effect of operating on the points $A\begin{pmatrix}2\\0\end{pmatrix}$, $B\begin{pmatrix}2\\1\end{pmatrix}$, $C\begin{pmatrix}0\\1\end{pmatrix}$ and $D\begin{pmatrix}0\\0\end{pmatrix}$ with the matrix $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. If these points become A', B', C' and D', investigate the effect of the matrix $\mathbf{N} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ on them. What do you deduce about the product **NM**?

(5) Draw diagrams to show the operating effect upon the points $\begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$ and $\begin{pmatrix}1\\1\end{pmatrix}$ of the matrices:

(*a*) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$.

(*Note*: For a situation like this it is useful to have a convention for distinguishing between these two: a useful method is to have a 'window'.) We show the matrices $\mathbf{M}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{M}_2 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$:

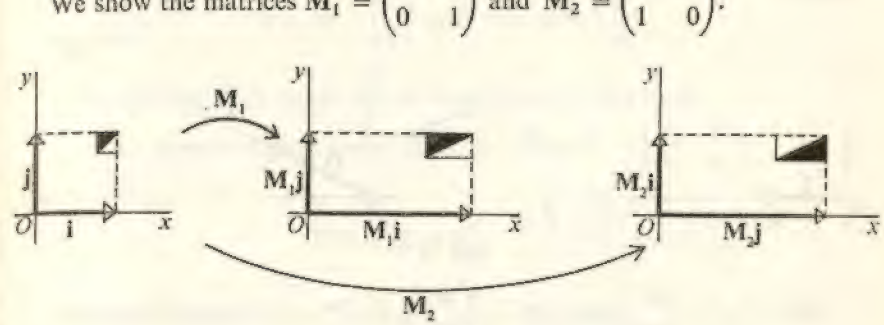


FIG. 52

The second mapping involves some 'turning over' process, i.e. reflection.

(6) Illustrate these matrices in a similar manner to question 1, and describe the mapping in words:

(*a*) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ (*d*) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ (*e*) $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

12.3 *Multiples and their transforms*

If the matrix is of the form $\begin{pmatrix} p & \cdot \\ q & \cdot \end{pmatrix}$, then not only is $\mathbf{Ti} = \begin{pmatrix}p\\q\end{pmatrix}$ but $\mathbf{T}(\lambda\mathbf{i}) = \begin{pmatrix} p & \cdot \\ q & \cdot \end{pmatrix}\begin{pmatrix}\lambda\\0\end{pmatrix} = \begin{pmatrix}\lambda p\\\lambda q\end{pmatrix}$. Thus, $\mathbf{T}(\lambda\mathbf{i}) = \lambda\mathbf{Ti}$.

In words, the transform of the multiple is the multiple of the transform, or *transformation* and *scalar multiplication* are commutative upon **i** and similarly upon **j**.

On our graph paper, if we make equal intervals upon the x-axis (parallel to **i**) as O, A, B, C, D, etc., then the transforms O', A', B', C', D', etc. are at equal intervals along another straight line. (Notice that for every 2×2 matrix the transform of O is still O, i.e. our O' coincides with O.)

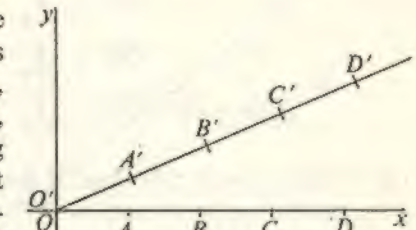


FIG. 53

12.4 *Sums and their transforms*

We have seen that $\mathbf{T}(\lambda\mathbf{i}) = \lambda\mathbf{Ti}$. This can be shown diagrammatically thus:

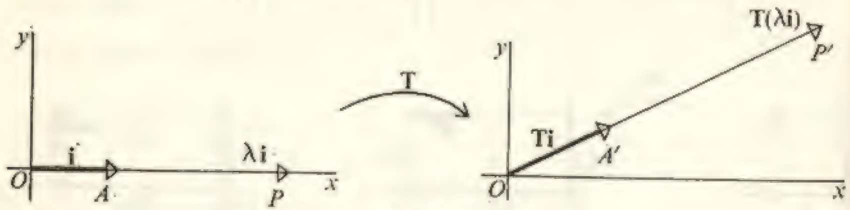


FIG. 54

Since $OA' = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $OP' = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \end{pmatrix}$, $OA'P'$ is a straight line; and, moreover, the ratio $\dfrac{OP'}{OA'} = \lambda = \dfrac{OP}{OA}$.

A similar result holds good for $\mathbf{j}$ and $\mu\mathbf{j}$, and we can incorporate the two results on one figure.

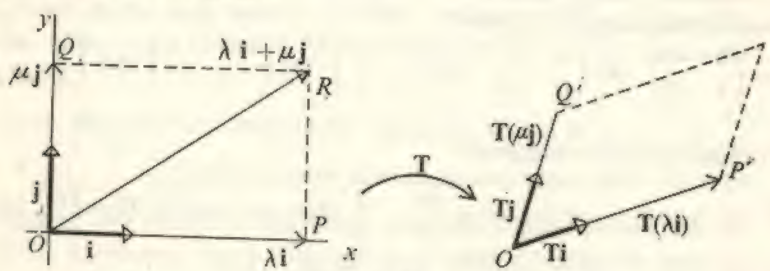


FIG. 55

We have developed the figure to show in each case the sum of the position vectors $\lambda\mathbf{i}$ and $\mu\mathbf{j}$ before and after their transformation:

$$\begin{array}{ccc} \overline{OP} & \xrightarrow{\mathbf{T}} & \overline{O'P'} \\ \overline{OQ} & \xrightarrow{\mathbf{T}} & \overline{O'Q'} \\ \overline{OR} = \overline{OP}+\overline{OQ} & \xrightarrow{\mathbf{T}} & ? \end{array}$$

It would be a very tidy result if the transform of R were the diagonal of the parallelogram formed by $O'P'$ and $O'Q'$; that is, if *the transform of a sum is the sum of its transforms*, i.e. if we could prove a theorem of the form $\mathbf{T}(\mathbf{u}+\mathbf{v}) = \mathbf{Tu}+\mathbf{Tv}$.

12.5 We shall not attempt to prove it in general—yet—but only for a case where $\mathbf{u}$ and $\mathbf{v}$ are multiples of the base vectors $\mathbf{i}$ and $\mathbf{j}$; i.e.

$$\mathbf{u} = \lambda\mathbf{i} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \text{ and } \mathbf{v} = \mu\mathbf{j} = \begin{pmatrix} 0 \\ \mu \end{pmatrix}.$$

In fact, we shall prove first that $\mathbf{T}(\lambda\mathbf{i}+\mu\mathbf{j}) = \lambda\mathbf{Ti}+\mu\mathbf{Tj}$.

For a general matrix transformation, where $\mathbf{T} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$,

$$\begin{aligned} \mathbf{T}(\lambda\mathbf{i}+\mu\mathbf{j}) &= \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\left(\lambda\begin{pmatrix} 1 \\ 0 \end{pmatrix}+\mu\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &= \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \\ &= \begin{pmatrix} \lambda a_1+\mu b_1 \\ \lambda a_2+\mu b_2 \end{pmatrix} \\ &= \lambda\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \mu\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \lambda\mathbf{Ti}+\mu\mathbf{Tj} \end{aligned}$$

and the theorem is proved.

This shows us how to find the transform of any vector very simply, by using the images of $\mathbf{i}$ and $\mathbf{j}$ and building up parallelogram-wise. The diagram shows $\mathbf{T}(3\mathbf{i}+2\mathbf{j})$ derived by adding $3\mathbf{Ti}+2\mathbf{Tj}$:

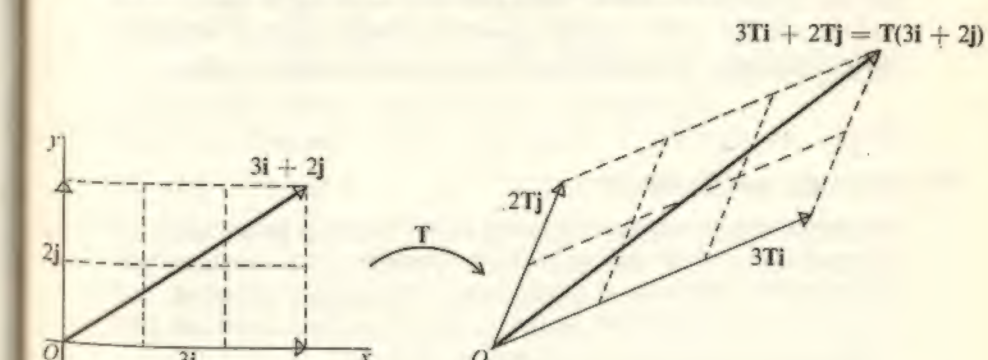


FIG. 56

12.6 The theorem is also true in the general case, i.e. when $\mathbf{u}$ and $\mathbf{v}$ are *not* multiples of the basis vectors, that $\mathbf{T}(p\mathbf{u}+q\mathbf{v}) = p\mathbf{Tu}+q\mathbf{Tv}$.†

For if $\mathbf{u} = x_1\mathbf{i}+y_1\mathbf{j}$ and $\mathbf{v} = x_2\mathbf{i}+y_2\mathbf{j}$, the left-hand side is

$$\begin{aligned}
\mathbf{T}(p\mathbf{u}+q\mathbf{v}) &= \mathbf{T}(px_1\mathbf{i}+py_1\mathbf{j}+qx_2\mathbf{i}+qy_2\mathbf{j}) \\
&= \mathbf{T}[(px_1+qx_2)\mathbf{i}+(py_1+qy_2)\mathbf{j}] \\
&= (px_1+qx_2)\mathbf{Ti}+(py_1+qy_2)\mathbf{Tj} \text{ (just proved)} \\
&= px_1\mathbf{Ti}+py_1\mathbf{Tj}+qx_2\mathbf{Ti}+qy_2\mathbf{Tj} \\
&= p(x_1\mathbf{Ti}+y_1\mathbf{Tj})+q(x_2\mathbf{Ti}+y_2\mathbf{Tj}) \\
&= p\mathbf{T}(x_1\mathbf{i}+y_1\mathbf{j})+q\mathbf{T}(x_2\mathbf{i}+y_2\mathbf{j}) \text{ (just proved)} \\
&= p\mathbf{Tu}+q\mathbf{Tv}
\end{aligned}$$

and the theorem is proved in general.

Thus a matrix is an operating mechanism which performs linear mappings, and, as we saw in the last example of exercise 11, any linear mapping may be written in the form of a matrix multiplying a vector. So matrix operations and linear mappings of vectors are merely different manifestations of the same situation, and each might be illuminated by a study of the other.

12.7 One final piece of work is needed to tidy up the connection between linear mappings and matrices. When two linear mappings **R** and **S** follow one another (**R** first) there is a composite mapping embodying the two called **SR**. When two matrices **M** and **N** operate one after the other upon a vector we have a product matrix **NM**. But the product rule for matrices has already been laid down—we ought now to show that the rule for matrix multiplication is a natural result of successive linear mappings. The reader is referred to the appendix for this.

12.8 *Associative rule for matrices*

The connection between matrices and linear mappings immediately has its reward, for we now need to prove that matrices are associative under multiplication. But since a matrix is a linear mapping operation, and we have already proved that mappings are associative, there is no need to prove separately that matrices are associative.

† The student may accept this statement at a first reading and omit the proof.

12.9 *Example*

Rotation through a given angle anti-clockwise. If $\mathbf{T}_\alpha$ is the matrix for a rotation through an angle α,
$\mathbf{T}_\alpha\mathbf{i} = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}$ and $\mathbf{T}_\alpha\mathbf{j} = \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}$;
and since these give the columns of our transformation matrix, this must be $\mathbf{T}_\alpha = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$.

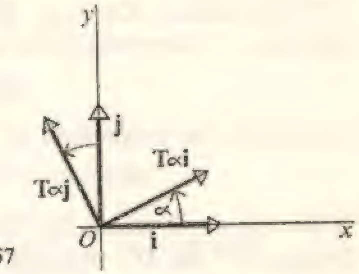


FIG. 57

The geometry indicates that this is a linear mapping; we also know from the geometry that to follow with a rotation of β would mean a total rotation of $\alpha+\beta$. This means $\mathbf{T}_{\alpha+\beta} = \mathbf{T}_\beta\mathbf{T}_\alpha$

$$\Leftrightarrow \begin{pmatrix}\cos\overline{\alpha+\beta} & -\sin\overline{\alpha+\beta}\\ \sin\overline{\alpha+\beta} & \cos\overline{\alpha+\beta}\end{pmatrix} = \begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$$

$$= \begin{pmatrix}\cos\beta\cos\alpha-\sin\beta\sin\alpha & -(\sin\beta\cos\alpha+\cos\beta\sin\alpha)\\ \sin\beta\cos\alpha+\cos\beta\sin\alpha & \cos\beta\cos\alpha-\sin\beta\sin\alpha\end{pmatrix}$$

$$\Leftrightarrow \begin{aligned}\cos\overline{\alpha+\beta} &= \cos\alpha\cos\beta-\sin\alpha\sin\beta\\ \sin\overline{\alpha+\beta} &= \sin\alpha\cos\beta+\cos\alpha\sin\beta.\end{aligned}$$

EXERCISE 12b

(1) Draw diagrams of the unit square to illustrate the operation of the matrices $\mathbf{X} = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$ and $\mathbf{Y} = \begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix}$.

Can you deduce by looking at these illustrations what the products **XY** and **YX** will be? Verify by direct multiplication.

(2) Draw diagrams of the unit square to illustrate the matrices $\mathbf{X} = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$ and $\mathbf{Y} = \begin{pmatrix}1 & -0\\ -0 & 3\end{pmatrix}$. Can you predict the value of the product **XY** by looking at the mappings? Verify by multiplication.

(3) Write down the matrix **R** which represents a reflection in the line $y = x\tan\frac{1}{2}\alpha$. Evaluate $\mathbf{R}^2$ and deduce that $\cos^2\alpha+\sin^2\alpha = 1$.

(4) Find the matrix $\mathbf{R}_\alpha$ which represents a reflection in the line $y = x\tan\frac{1}{2}\alpha$. Show that, if $\mathbf{R}_\beta$ represents a reflection in the line $y = x\tan\frac{1}{2}\beta$, $\mathbf{R}_\beta\mathbf{R}_\alpha$ represents a rotation, and find the angle of rotation.

(5) A matrix which has the property that the sum of the elements in each row is unity is called a *stochastic* matrix. Prove that the product of two such 2×2 matrices is also stochastic.

If the columns also have the same property the matrix is doubly stochastic. Is the product of two such 2×2 matrices doubly stochastic?

(6) Writing the polynomial ax^3+bx^2+cx+d as a 4-vector $\begin{pmatrix} a\\ b\\ c\\ d\end{pmatrix}$, what is the matrix $\mathbf{D}$ which has the effect of differentiating this vector? Evaluate $\mathbf{D}^2$ and $\mathbf{D}^4$, verifying that $\mathbf{D}^4 = \mathbf{0}$.

(7) Show that $\begin{pmatrix} 1 & 0\\ c/a & 1\end{pmatrix}\begin{pmatrix} 1 & ab/\Delta\\ 0 & 1\end{pmatrix}\begin{pmatrix} a & 0\\ 0 & \Delta/a\end{pmatrix} = \begin{pmatrix} a & b\\ c & d\end{pmatrix}$ where $\Delta = ad-bc$.

Investigate the geometric transformations of each of the three matrices on the left-hand side, and show that only one of them changes area. Hence show that the matrix $\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ increases area by a factor $ad-bc$.

(8) Show that the matrix $\mathbf{M} = \begin{pmatrix} 4 & -3\\ -1 & 2\end{pmatrix}$ operating on *any* vector in the direction $\begin{pmatrix} 1\\ 1\end{pmatrix}$ leaves it unaltered. Show also that there are vectors in the direction $\begin{pmatrix} +3\\ -1\end{pmatrix}$ whose direction is unaltered by the matrix, but that these are enlarged by a factor 5. Investigate the action of $\mathbf{M}^2$ upon these vectors.

chapter 13

MORE ABOUT LINEAR EQUATIONS

13.1 In chapter 5 we considered the set of linear equations:

$$\begin{aligned} a_1x+b_1y+c_1z &= d_1\\ a_2x+b_2y+c_2z &= d_2 \qquad (\mathbf{d}\neq 0)\\ a_3x+b_3y+c_3z &= d_3\end{aligned}$$

as the solution in x, y and z of the vector equation

$$x\begin{pmatrix} a_1\\ a_2\\ a_3\end{pmatrix} + y\begin{pmatrix} b_1\\ b_2\\ b_3\end{pmatrix} + z\begin{pmatrix} c_1\\ c_2\\ c_3\end{pmatrix} = \begin{pmatrix} d_1\\ d_2\\ d_3\end{pmatrix}, \qquad (\mathbf{d}\neq 0).$$

By considering certain special cases we decided that, provided the coefficients of the equations, i.e. the elements in the matrix $\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix}$ possessed certain numerical properties, the equations would have a unique solution. This property was investigated only for an echelon set of equations (i.e. a_2, a_3 and $b_3 = 0$) when we stated that the value of the determinant was $a_1b_2c_3$.

In the light of our subsequent work we can regard the matrix as a mapping agency; acting on any 3-vector $\mathbf{v} = \begin{pmatrix} x\\ y\\ z\end{pmatrix}$ it transforms it into an image vector, which we can write as $\mathbf{Mv}$ or as $\begin{pmatrix} x'\\ y'\\ z'\end{pmatrix}$.

We now regard the problem in a new light; given a non-zero vector $\begin{pmatrix} d_1\\ d_2\\ d_3\end{pmatrix}$, can we find a unique original of which this is the image, i.e. *is the transformation reversible*? Given the right-hand vector, can we find the $\begin{pmatrix} x\\ y\\ z\end{pmatrix}$ which corresponds to it?

We give some useful technical terms in this connection. A set of equations which can be solved uniquely for a r.h. vector non-zero is called *regular*; and the l.h. matrix (if it is square) is called *non-singular*.

We shall consider some simple cases in the next two sections, before attempting a more general treatment.

13.2 Let us consider the case (which we know to be regular):

$$\begin{aligned} 5x+7y &= x' \\ 2x+3y &= y' \end{aligned} \quad \text{or} \quad \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Direct solution gives us:

$$\begin{aligned} x &= 3x'-7y' \\ y &= -2x'+5y' \end{aligned} \quad \text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

That this is the reverse mapping may easily be verified by direct substitution for x and y in the equations $\begin{aligned} 5x+7y &= x' \\ 2x+3y &= y'. \end{aligned}$

We have:

$$\begin{aligned} &5(3x'-7y')+7(-2x'+5y') \\ &= 15x'-35y'-14x'+35y' \\ &= x' \end{aligned}$$

as required. Similarly we find that:

$$\begin{aligned} &2(3x'-7y')+3(-2x'+5y') \\ &= 6x'-14y'-6x'+15y' \\ &= y'. \end{aligned}$$

It is interesting to see what happens with the matrix form. Substituting for

$$\begin{pmatrix} x \\ y \end{pmatrix} \text{ in } \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \text{ we have}$$

$$\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

We notice that the product of $\begin{pmatrix} 5 & 7 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 3 & -7 \\ -2 & +5 \end{pmatrix}$ is the identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Thus $\begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}$ is said to be the *inverse* matrix† of $\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$ and represents the inverse mapping.

† This is an example of the general algebraic term. See chapters 1 and 6. Strictly we should say it is the inverse for matrix multiplication.

It is also apparent that $\begin{aligned} x' &= 5x+7y \\ y' &= 2x+3y \end{aligned}$ is the inverse mapping of $\begin{aligned} x &= 3x'-7y' \\ y &= -2x'+5y' \end{aligned}$, so that we should also expect $\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$ to be the inverse of $\begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}$. Direct multiplication shows this to be true:

$$\begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

13.3 We shall now consider a triad of equations, which we also know to be regular. (How?)

$$\begin{aligned} x+py+qz &= x' \\ y+rz &= y' \\ z &= z' \end{aligned} \quad \text{or} \quad \begin{vmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{vmatrix}\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x' \\ y' \\ z' \end{vmatrix}$$

Direct solution gives us:

$$\begin{aligned} x &= x'-py'+(pr-q)z' \\ y &= y'-rz' \\ z &= z' \end{aligned} \quad \text{or} \quad \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} 1 & -p & pr-q \\ 0 & 1 & -r \\ 0 & 0 & 1 \end{vmatrix}\begin{vmatrix} x' \\ y' \\ z' \end{vmatrix}.$$

Once again the product of the matrices is of interest:

$$\begin{vmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{vmatrix}\begin{vmatrix} 1 & -p & pr-q \\ 0 & 1 & -r \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -p+p & pr-q-pr+q \\ 0 & 1 & -r+r \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

and

$$\begin{vmatrix} 1 & -p & pr-q \\ 0 & 1 & -r \\ 0 & 0 & 1 \end{vmatrix}\begin{vmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & p-p & q-pr+pr-q \\ 0 & 1 & r-r \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

Again we say that:

$$\begin{vmatrix} 1 & -p & pr-q \\ 0 & 1 & -r \\ 0 & 0 & 1 \end{vmatrix} \text{ is the } \textit{inverse} \text{ of } \begin{vmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{vmatrix}$$

and vice versa. The notation used is the usual one—the inverse of a matrix $\mathbf{M}$ is written $\mathbf{M}^{-1}$ and we have $\mathbf{M}\mathbf{M}^{-1} = \mathbf{M}^{-1}\mathbf{M} = \mathbf{I}$, the identity or *unit* matrix of the appropriate size $n \times n$.

Let us now use the machinery outlined above to find the inverse of the matrix $\begin{vmatrix} a & l & m \\ 0 & b & n \\ 0 & 0 & c \end{vmatrix}$.

We consider the equations
$$\begin{aligned} ax+ly+mz &= x' \\ by+nz &= y'. \\ cz &= z' \end{aligned}$$

Then $z = \dfrac{z'}{c}$, provided $c \neq 0$

$$by = y' - \frac{nz'}{c} \iff y = \frac{y'}{b} - \frac{nz'}{bc} \qquad (b \neq 0)$$

$$ax = x' - \frac{ly'}{b} + \frac{lnz'}{bc} - \frac{mz'}{c}$$

$$\iff x = \frac{x'}{a} - \frac{ly'}{ab} + \left(\frac{ln}{abc} - \frac{m}{ac}\right) z' \qquad (a \neq 0).$$

Then
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} \frac{1}{a} & -\frac{l}{ab} & \frac{ln}{abc} - \frac{m}{ac} \\ 0 & \frac{1}{b} & -\frac{n}{bc} \\ 0 & 0 & \frac{1}{c} \end{bmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \qquad (a, b, c \neq 0).$$

The inverse of $\begin{pmatrix} a & l & m \\ 0 & b & n \\ 0 & 0 & c \end{pmatrix}$ is $\begin{bmatrix} \frac{1}{a} & -\frac{l}{ab} & \frac{ln}{abc} - \frac{m}{ac} \\ 0 & \frac{1}{b} & -\frac{n}{bc} \\ 0 & 0 & \frac{1}{c} \end{bmatrix}$ provided none of a, b and $c = 0$ or, equivalently, $abc \neq 0$.

Direct multiplication gives us:

$$\begin{pmatrix} a & l & m \\ 0 & b & n \\ 0 & 0 & c \end{pmatrix} \begin{bmatrix} \frac{1}{a} & -\frac{l}{ab} & \frac{ln}{abc} - \frac{m}{ac} \\ 0 & \frac{1}{b} & -\frac{n}{bc} \\ 0 & 0 & \frac{1}{c} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{l}{b} + \frac{l}{b} & \frac{ln}{bc} - \frac{m}{c} - \frac{ln}{bc} + \frac{m}{c} \\ 0 & 1 & -\frac{n}{c} + \frac{n}{c} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

the unit matrix of order 3.

The examples given illustrate the following logical situation:

(i) Existence of solution $\iff$ The equations are regular (i.e. by definition matrix is non-singular).

(*ii*) Existence of solution $\iff$ Inverse matrix exists.

Hence every non-singular matrix has an inverse.

EXERCISE 13a

(1) Given that $\left.\begin{aligned} 2x_1+5x_2 &= X_1 \\ 3x_2 &= X_2 \end{aligned}\right\}$ express x_1, x_2 in terms of X_1, X_2.

Express your results in matrix form, and show that the products of the two 2×2 matrices *both ways round* are $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(2) Given that $\mathbf{Mu} = \mathbf{v}$,

where $\qquad \mathbf{M} = \begin{pmatrix} 1 & 4 & 5 \\ 0 & 1 & 7 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ $\qquad \ldots (1)$

write the data in equation form, solve for x, y, z in terms of X, Y, Z and write the result in the form

$$\mathbf{u} = \mathbf{Nv} \qquad \ldots (2)$$

Verify by direct multiplication that $\mathbf{MN} = \mathbf{NM} = \mathbf{I}$. Using these results, state short matrix proofs that (1) $\iff$ (2) and (2) $\iff$ (1). Make clear at which points you use the associative law for matrix multiplication.

(3) By re-writing the equations $\left\{\begin{aligned} 2x+3y &= X \\ x+2y &= Y \end{aligned}\right.$ to give x and y in terms of X and Y, find the inverse of the matrix $\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$.

(4) Show that the inverse of an echelon matrix of form $\begin{pmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{pmatrix}$ is of similar form.

Can you give another type of triangular matrix (i.e. with zeros in one corner) for which this is true? An example in the text shows that it is true for any echelon matrix if none of the main diagonal coefficients is zero; but you are asked to consider also other triangular matrices, i.e. in which the zeros are not in the *left-hand bottom* corner. (Echelon matrices form a subset of triangular matrices.) It will be sufficient to

consider matrices in which a diagonal of coefficients consists of 1's. Try 2×2 matrices first.

(5) By solving the equations $\begin{aligned} 2x+y+z &= X, \\ x+y+z &= Y, \\ -x \quad +z &= Z, \end{aligned}$ find the inverse of the matrix $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$. Check your answer.

(6) Use the method of question 5 to find the inverses of the matrices:

$$(a) \begin{pmatrix} -1 & 4 & 1 \\ -2 & -2 & -1 \\ 0 & 3 & 1 \end{pmatrix} \qquad (b) \begin{pmatrix} 1 & -1 & 4 \\ 0 & 1 & -4 \\ 2 & -3 & 13 \end{pmatrix}$$

(*Note*: This is not usually the best way of finding the inverse matrix, but it is reasonably efficient.)

13.4 We shall now show a method of dealing with *any* triad of equations which quickly shows whether they are regular or not, and leads to a solution if they are.

Choleski's method for solving linear equations is an ingenious method which relies on factorising the operating matrix into two matrices which, owing to their form, are called triangular. We illustrate with the set of linear equations $\begin{pmatrix} 1 & 3 & 1 \\ 2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \\ 8 \end{pmatrix}$.

We factorise the matrix $\begin{pmatrix} 1 & 3 & 1 \\ -2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix}$ into the form

$$\begin{pmatrix} 1 & 3 & 1 \\ -2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix} = \begin{pmatrix} . & 0 & 0 \\ . & . & 0 \\ . & . & . \end{pmatrix}\begin{pmatrix} . & . & . \\ 0 & . & . \\ 0 & 0 & . \end{pmatrix} = \mathbf{L}\,.\,\mathbf{U} \text{ (say)},$$

where the dots are to be determined, and the triangular matrices are called **L** (lower triangular) and **U** (upper triangular).

The reader is advised to try this, but a few trials are enough to show that the factorisation is by no means unique.† We resolve this by

† The matrix product involves 9 facts and we have 12 unknowns.

making each of the diagonal elements of **U** equal to unity. Hence:

$$\begin{pmatrix} 1 & 3 & 1 \\ -2 & -3 & 4 \\ 1 & 5 & 2 \end{pmatrix} = \begin{pmatrix} . & 0 & 0 \\ . & . & 0 \\ . & . & . \end{pmatrix}\begin{pmatrix} 1 & . & . \\ 0 & 1 & . \\ 0 & 0 & 1 \end{pmatrix}$$

and the complete factorisation is now uniquely possible. We find:

$$\begin{pmatrix} 1 & 3 & 1 \\ -2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 0 \\ 1 & 2 & -7 \end{pmatrix}\begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

and the original equations become:

$$\mathbf{L}\,.\,\mathbf{U}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \\ 8 \end{pmatrix}.$$

We now put $\mathbf{U}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and the equations become:

$$\mathbf{L}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \\ 8 \end{pmatrix} \text{ i.e. } \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 0 \\ 1 & 2 & -7 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \\ 8 \end{pmatrix}.$$

This equation is directly soluble for a, b, c, for on re-writing we have:

$$\left.\begin{aligned} a &= 3 \\ -2a+3b &= -9 \\ a+2b-7c &= 8 \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} a &= 3 \\ b &= -1. \\ c &= -1 \end{aligned}\right.$$

Now, since $\mathbf{U}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, we have

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix};$$

or, re-writing, $\left.\begin{aligned} x+3y+z &= 3 \\ y+2z &= -1 \\ z &= -1 \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} x &= 1 \\ y &= 1 \\ z &= -1, \end{aligned}\right.$

obtained in the order $z = -1$, $y = 1$, $x = 1$.

From our previous work, we know that an echelon set of equations is bound to be regular if the product of the elements on the diagonal is

non-zero. For this set we have split the matrix into two triangular matrices **L** and **U**, where we have stipulated that the diagonal elements of **U** should each be unity. It is therefore the diagonal elements of **L** which are critical: if their product is zero the equations will have either no unique solution or no solution at all. This product then determines whether the set of equations is regular or not, and vanishes (or does not vanish) with the determinant itself.

In practice, the writing out of this process is much quicker than appears above. We give two examples, one regular and one non-regular.

13.5 *Worked examples*

(*a*) Solve the linear equations:

$$\begin{aligned} 2x+6y+14z &= 6\\ 4x+9y+13z &= 3\\ -x+3y+24z &= 17. \end{aligned}$$

Writing these equations $\mathbf{M}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}6\\3\\17\end{pmatrix}$ and factorising **M** gives:

$$\begin{pmatrix}2 & 6 & 14\\4 & 9 & 13\\-1 & 3 & 24\end{pmatrix} = \begin{pmatrix}2 & 0 & 0\\4 & -3 & 0\\-1 & 6 & 1\end{pmatrix}\begin{pmatrix}1 & 3 & 7\\0 & 1 & 5\\0 & 0 & 1\end{pmatrix} = \mathbf{LU}.$$

Since the product of the diagonal elements of **L** is -6, i.e. non-zero,

we proceed by putting $\mathbf{U}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}$

$$\Leftrightarrow \quad \begin{pmatrix}2 & 0 & 0\\4 & -3 & 0\\-1 & 6 & 1\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = \begin{pmatrix}6\\3\\17\end{pmatrix}$$

$$\Leftrightarrow \quad a = 3,\ b = 3,\ c = 1$$

$$\Leftrightarrow \quad \begin{pmatrix}1 & 3 & 7\\0 & 1 & 5\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}3\\3\\1\end{pmatrix}$$

$$\Leftrightarrow \quad z = 1,\ y = -2,\ x = 2.$$

(*b*) Solve the linear equations:

$$\begin{aligned} x+2y-\ z &= -1\\ 2x+7y-5z &= 3\\ 7x+6y+\ z &= -6. \end{aligned}$$

Writing these equations as $\mathbf{M}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}-1\\3\\-6\end{pmatrix}$ and factorising **M** gives:

$$\begin{pmatrix}1 & 2 & -1\\2 & 7 & -5\\7 & 6 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\2 & 3 & 0\\7 & -8 & 0\end{pmatrix}\begin{pmatrix}1 & 2 & -1\\0 & 1 & -1\\0 & 0 & 1\end{pmatrix}.$$

Examination of **L** shows that the product of the diagonal elements is zero, i.e. there is no *unique* solution.

Proceeding, for there may be solutions, by writing $\mathbf{U}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}$

$$\Leftrightarrow \quad \begin{pmatrix}1 & 0 & 0\\2 & 3 & 0\\7 & -8 & 0\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = \begin{pmatrix}-1\\3\\-6\end{pmatrix}$$

$$\Leftrightarrow \quad \begin{aligned} a &= -1\\ 2a+3b &= 3\\ 7a-8b &= -6. \end{aligned}$$

The first two equations tell us $a = -1$, $b = \frac{5}{3}$, but the third equation is not satisfied. Hence there are *no* solutions.

EXERCISE 13b

(1) Express $\mathbf{M} = \begin{pmatrix}1 & 2 & -1\\1 & 1 & 0\\4 & 8 & -1\end{pmatrix}$ in the form **LU** where $\mathbf{L} = \begin{pmatrix}a & 0 & 0\\b & c & 0\\d & e & f\end{pmatrix}$ and $\mathbf{U} = \begin{pmatrix}1 & g & h\\0 & 1 & k\\0 & 0 & 1\end{pmatrix}$.

If $\mathbf{L}\begin{pmatrix}X\\Y\\Z\end{pmatrix} = \begin{pmatrix}6\\5\\33\end{pmatrix}$, solve for X, Y, Z and proceed to solve the equation $\mathbf{U}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}X\\Y\\Z\end{pmatrix}$.

Show by direct substitution of values that

$$\begin{aligned} x+2y-z &= 6 \\ x+y \quad &= 5 \\ 4x+8y-z &= 33. \end{aligned}$$

(This example shows that the method of Choleski, so useful in computing, is not to be preferred to direct elimination in *simple* cases.)

(2) Show, using the associative law, that if $\mathbf{A} = \mathbf{LU}$ then the inverse $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$ (*N.B.* showing that $(\mathbf{U}^{-1}\mathbf{L}^{-1})(\mathbf{LU}) = \mathbf{I}$ is only half the answer.)

Obtain the inverses of **L** and **U** in question 1, and then that of **M**.

Hence solve immediately, since $\mathbf{Mu} = \mathbf{v} \Leftrightarrow \mathbf{u} = \mathbf{M}^{-1}\mathbf{v}$, the triad of equations:

$$\begin{aligned} 2x+4y-2z &= 12 \\ x+y \quad &= 5 \\ 4x+8y-z &= 27. \end{aligned}$$

(3) Given $\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 \\ 2 & -5 & -4 \\ 3 & -6 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 29 & -2 & -8 \\ 14 & -1 & -4 \\ -3 & 0 & 1 \end{pmatrix}$, evaluate **AB**.

Make use of your result to solve:

(*a*) $\begin{aligned} x-2y \quad &= 3 \\ 2x-5y-4z &= 5 \\ 3x-6y+z &= 9 \end{aligned}$ (*b*) $\begin{aligned} 29x-2y-8z &= -3 \\ 14x-y-4z &= 1 \\ -3x \quad +z &= -1 \end{aligned}$

(4) Verify that if $\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 0 & 2 \end{pmatrix}$, then $\mathbf{L}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -\frac{3}{2} & 0 & \frac{1}{2} \end{pmatrix}$.

Write any simple numerical matrix of the form $\mathbf{U} = \begin{pmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{pmatrix}$, and hence derive two matrices **LU**, $\mathbf{U}^{-1}\mathbf{L}^{-1}$ which are inverses. Verify this result directly.

(5) Solve the following sets of equations by using Choleski's method:

(*a*) $\begin{aligned} x+2y+6z &= 0, \\ 3x+8y+12z &= 10, \\ 4x+7y+8z &= 14. \end{aligned}$ (*b*) $\begin{aligned} 3x+12y-9z &= 0, \\ 2x+4y+2z &= 6, \\ -x+2y+3z &= 3. \end{aligned}$

(*c*) $\begin{aligned} 2x+8y-4z &= 6, \\ x+5y+z &= 7, \\ 3x+11y-8z &= 6. \end{aligned}$ (*d*) $\begin{aligned} 3x-3y+3z &= 8, \\ 4x-3y+6z &= 15, \\ 2x-3y+z &= 3. \end{aligned}$

chapter 14

DETERMINANTS

14.1 We have already seen that the determinant of the triangular matrix $\begin{pmatrix} a_1 & p & q \\ 0 & b_2 & r \\ 0 & 0 & c_3 \end{pmatrix}$ is $a_1b_2c_3$, and that the non-vanishing of this is the condition for the set of equations

$$\begin{pmatrix} a_1 & p & q \\ 0 & b_2 & r \\ 0 & 0 & c_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \qquad (\mathbf{d} \neq 0)$$

to have a unique solution. We now generalise this result and remind the reader that a determinant is a pure number derived from all the coefficients of a matrix, and for the matrix $\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ it is written symbolically as:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \det \mathbf{M}.$$

The value of det **M** we shall now affirm to be

$$= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$

At the moment we have baldly stated this expression and it remains to establish that it has the properties we require of it. But before we do this we investigate the properties of the expression.

The expression for Δ is best remembered by breaking it up, e.g.

$$\begin{aligned} \Delta &= a_1(b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2) \\ &= a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}. \end{aligned}$$

The connection between the elements a_1, b_1 and c_1 and the small determinants which multiply them is simple. We take each element of the top row in turn, and cross out the row and column containing that element; elements which remain as seen in each case make up the 2×2 multiplying determinants:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

giving the required products:

$$a_1 \times \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}, \quad b_1 \times \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}, \quad c_1 \times \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$

When the three terms of the expansion are obtained in this way, it only remains to remember that the sign of the middle term is negative.

This method is called the 'expansion of the 3×3 determinant by the first row'. It is particularly quick if one of the first-row elements happens to be zero, e.g.

$$\begin{vmatrix} 2 & 3 & 0 \\ 5 & 4 & 7 \\ 8 & -1 & 6 \end{vmatrix} = 2\times\begin{vmatrix} 4 & 7 \\ -1 & 6 \end{vmatrix} - 3\times\begin{vmatrix} 5 & 7 \\ 8 & 6 \end{vmatrix}$$

$$= 2\times 31 - 3\times(-26)$$
$$= 140.$$

EXERCISE 14a

Evaluate the following determinants by the first row:

(1) (*a*) $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 6 & 1 & 1 \end{vmatrix}$ (*b*) $\begin{vmatrix} 1 & 1 & 0 \\ 2 & 3 & 6 \\ 1 & 1 & -2 \end{vmatrix}$

(*c*) $\begin{vmatrix} 2 & 1 & 3 \\ 3 & 4 & -2 \\ 5 & 5 & 1 \end{vmatrix}$ (*d*) $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$

(2) Expand the determinants given, noticing the relation between the rows of the determinants and the final values. Draw any conclusion you can.

(*i*) $\begin{vmatrix} 3 & 1 & 7 \\ 2 & -1 & 3 \\ 0 & -1 & 2 \end{vmatrix}$ and $\begin{vmatrix} 2 & -1 & 3 \\ 3 & 1 & 7 \\ 0 & -1 & 2 \end{vmatrix}$

(*ii*) $\begin{vmatrix} 1 & -1 & 3 \\ 1 & 4 & 7 \\ 2 & 6 & -1 \end{vmatrix}$ and $\begin{vmatrix} 2 & 6 & -1 \\ 1 & 4 & 7 \\ 1 & -1 & 3 \end{vmatrix}$

(*iii*) $\begin{vmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 3 & 3 & 3 \end{vmatrix}$ and $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ -1 & 1 & 4 \end{vmatrix}$

(*iv*) $\begin{vmatrix} 1 & -1 & 4 \\ 2 & 0 & 3 \\ 1 & 4 & 2 \end{vmatrix}$ and $\begin{vmatrix} 2 & -2 & 8 \\ 2 & 0 & 3 \\ 1 & 4 & 2 \end{vmatrix}$

(*v*) $\begin{vmatrix} 3 & 1 & -1 \\ 4 & 1 & -2 \\ 1 & 0 & 6 \end{vmatrix}$ and $\begin{vmatrix} 3 & 1 & -1 \\ 12 & 3 & -6 \\ 1 & 0 & 6 \end{vmatrix}$

(3) When the expression for Δ was collected up for the expansion by the top row, the terms a_1, b_1 and c_1 were isolated and their multipliers were found. Follow this working through for the third row and decide on a similar scheme for expanding by using it: repeat this process for the second row.

(4) Expand these determinants using whichever row you think most suitable.

(*i*) $\begin{vmatrix} 1 & 2 & 6 \\ 0 & 3 & 1 \\ 2 & 0 & 0 \end{vmatrix}$ (*ii*) $\begin{vmatrix} 1 & -2 & 7 \\ 0 & 0 & 1 \\ 2 & 3 & -3 \end{vmatrix}$ (*iii*) $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 0 & 2 & -1 \end{vmatrix}$

(*iv*) $\begin{vmatrix} 3 & 2 & 1 \\ 1 & 4 & -5 \\ 0 & 3 & 0 \end{vmatrix}$ (*v*) $\begin{vmatrix} 3 & 1 & 0 \\ 2 & 3 & -1 \\ 6 & 4 & 2 \end{vmatrix}$ (*vi*) $\begin{vmatrix} 1 & 7 & 4 \\ 0 & 2 & 3 \\ -3 & 2 & 2 \end{vmatrix}$

(5) Compare the values of the determinants:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1+a_2 & b_1+b_2 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & b_1 \\ a_2+ka_1 & b_2+kb_1 \end{vmatrix},$$

and draw any conclusions you can about importing into one row a complete multiple of another row—for 2×2 determinants.

(6) Evaluate the following determinants, and by inspecting the rows, draw conclusions about the final values.

(*i*) $\begin{vmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1 & 6 & 1 \end{vmatrix}$ (*ii*) $\begin{vmatrix} 1 & 3 & 1 \\ 2+1 & 4+3 & -1+1 \\ 1 & 6 & 1 \end{vmatrix}$

(*iii*) $\begin{vmatrix} 1-1 & 3-6 & 1-1 \\ 2 & 4 & -1 \\ 1 & 6 & 1 \end{vmatrix}$ (*iv*) $\begin{vmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1-1 & 6-3 & 1-1 \end{vmatrix}$

(*v*) $\begin{vmatrix} 1 & 3 & 1 \\ 2-2\times 1 & 4-2\times 3 & -1-2\times 1 \\ 1 & 6 & 1 \end{vmatrix}$ (*vi*) $\begin{vmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 1+x & 6+3x & 1+x \end{vmatrix}$

14.2 *Evaluation of determinant by any row*

The reader will probably have developed a working rule for deciding how to evaluate a determinant by any row, but we shall summarise the rules here.

If we decide to use, say, the second row, we start off with the element a_2 and cross out the row and column containing a_2. This leaves a determinant,

$$\begin{vmatrix} \cancel{a_1} & b_1 & c_1 \\ \cancel{a_2} & \cancel{b_2} & \cancel{c_2} \\ \cancel{a_3} & b_3 & c_3 \end{vmatrix}$$

i.e. $\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}$, which is called the *minor determinant* for a_2 or, more simply, just plain *minor*.

We do the same thing for b_2 and c_2 and we have 3 terms

$$a_2 \times \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}, \quad b_2 \times \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} \quad \text{and} \quad c_2 \times \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}.$$

The only problem is which sign we should attach to the minor, and comparison with the expression Δ suggests

$$\Delta = -a_2 \times \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2 \times \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2 \times \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}.$$

Notice that a rule is emerging which also applies to the 2×2 case where the minor is the element in the opposite corner and

$$\begin{vmatrix} p & q \\ r & s \end{vmatrix} = p.s-q.r.$$

That is, the sign given to the minor depends on *how many moves you have to make from the top left-hand corner to reach your element*, e.g. if we have to expand using a_2, b_2 and c_2 then the signs are alternately $-+-$ in that order, and for a_3, b_3 and c_3 they will be $+-+$.

$$\text{Again: } \Delta = a_3\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} - b_3\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} + c_3\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Usually it is more convenient, rather than using the minors 'raw', to carry out at once any necessary change of sign: these rectified minors are called *cofactors*. It is common practice to use the symbol A_1 to denote the cofactor corresponding to a_1, and so on. Thus,

$$B_3 = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix},$$

and the expansions of the determinant are

$$a_1A_1+b_1B_1+c_1C_1$$
$$\text{or} \quad a_2A_2+b_2B_2+c_2C_2$$
$$\text{or} \quad a_3A_3+b_3B_3+c_3C_3.$$

Clearly this revised procedure has no advantage in numerical evaluations, but the tidiness of these results is a great help in theoretical work.

The method of evaluating 4×4 and larger determinants is a logical extension of the 3×3 method.

EXERCISE 14b

(1) Evaluate *all* the *cofactors* of the elements of the determinant

$$\begin{array}{ccc} (a_1) & (b_1) & (c_1) \end{array}$$
$$\begin{vmatrix} 2 & 3 & -1 \\ -5 & 8 & 4 \\ 3 & 0 & 2 \end{vmatrix}$$

For convenience in checking, set them out in the corresponding array

$$\begin{array}{ccc} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3. \end{array}$$

(2) Check that for values in question 1:

$$a_1A_1+b_1B_1+c_1C_1 = a_2A_2+b_2B_2+c_2C_2 \ (=\Delta);$$
$$a_1A_3+b_1B_3+c_1C_3 = 0;$$
$$B_1C_3-B_1C_3 = a_2\Delta.$$

(3) Given $\mathbf{M} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, $\mathbf{N} = \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix}$, find the product matrix **MN**.

Prove that $\det(\mathbf{MN}) = \det \mathbf{M} \times \det \mathbf{N}$.

(4) Given $\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 4 & 5 & 8 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & 0 \\ 1 & 1 & 0 \end{pmatrix}$, find the product matrices **AB** and **BA**. Show that although $\mathbf{AB} \neq \mathbf{BA}$, yet $\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A} \times \det \mathbf{B}$.

(5) Evaluate the cofactors of the determinant $\begin{vmatrix} 1 & 3 & 1 \\ 7 & 2 & 4 \\ 4 & -7 & 1 \end{vmatrix}$, laying them out in the manner described in question 1. What do you notice about the rows of this matrix of cofactors? What is the value of the original determinant?

(6) Treat the determinant $\begin{vmatrix} 2 & 3 & -1 \\ 4 & -1 & 6 \\ 0 & -7 & 8 \end{vmatrix}$ in the manner described in question 5.

(7) Evaluate the cofactors of the determinant $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix}$, and comment on your results.

14.3 *Properties of determinants*

In solving equations, and in manipulating the matrices of their coefficients, we used what we came to call *row operations*. We now look to see what these same row operations do to the value of the determinants of the coefficient matrix.

The row operations which were important were:

(*a*) Interchanging two rows.

(*b*) Adding a multiple of one row to another row, e.g. $r_3' = r_3 + \lambda r_2$.

(*c*) Multiplying all elements of a row by a number, in order, say, to clear fractions; e.g. $r_2' = \mu r_2$.

For equations with integral coefficients we sometimes perform a more general step such as $r_3' = \lambda r_3 + \mu r_2$. This is not permitted in numerical work on hand machines, and it is, anyway, a combination of methods 2 and 3, so we shall not consider it here.

(*a*) *Interchange of two rows*

We suppose first that we interchange two adjacent rows, and we shall take the first and second as our example.

The problem, then, is

$$\text{if } \Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \text{ what is the value of } \begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}?$$

Now $\Delta = a_1A_1 + b_1B_1 + c_1C_1$, but if we expand the second determinant by the second row, we find that, according to the rule given for the signs of minors, each cofactor, although numerically the same as before, has changed sign. Hence the value of the second determinant is $-\Delta$. The reader should show that this is true if we change over any two *adjacent* rows. If we now want to change *any* two rows we shall do it in the following way: we start with the first row to be moved, and gradually, changing always adjacent rows, we put it into the required position, taking, say, n moves. We then move the other row back step by step, and this takes *one less move*, i.e. $(n-1)$ moves. The total $n+(n-1) = 2n-1$ is necessarily odd, giving us an odd number of changes of sign. Hence interchanging *any* two rows changes the sign of the determinant.

Corollary: If two rows are identical $\Delta = 0$, for if we interchange the two identical rows we have $\Delta = -\Delta \Rightarrow \Delta = 0$.

(*b*) *Addition to one row of a multiple of another row*

We shall discuss this in relation to the first row only—the treatment is the same whichever row we take.

If we form a new determinant,

$$\Delta' = \begin{vmatrix} a_1+\lambda a_2 & b_1+\lambda b_2 & c_1+\lambda c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

and expand by the top row,

$$\begin{aligned} \Delta' &= (a_1+\lambda a_2)A_1 + (b_1+\lambda b_2)B_1 + (c_1+\lambda c_2)C_1 \\ &= a_1A_1 + b_1B_1 + c_1C_1 + \lambda(a_2A_1 + b_2B_1 + c_2C_1) \\ &= \Delta + \lambda E \text{ (say).} \end{aligned}$$

But $E = \begin{vmatrix} a_2 & b_2 & c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$ (two rows equal). Hence $\Delta' = \Delta$, and this row operation leaves Δ unchanged.

(c) *Multiplication of the elements of one row by a factor* μ

Again we consider the first row only and form

$$\Delta' = \begin{vmatrix} \mu a_1 & \mu b_1 & \mu c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Expanding by the top row, $\Delta' = \mu a_1 A_1 + \mu b_1 B_1 + \mu c_1 C_1$
$= \mu\Delta.$

These three properties of determinants enable us to ease the evaluation of a determinant quite considerably by carrying out judiciously chosen row operations. Here are two worked examples, in which a condensed notation is used to explain how a new row is formed from the old:

(a) To evaluate $\begin{vmatrix} 1 & 3 & 4 \\ 13 & 12 & 10 \\ 27 & 27 & 24 \end{vmatrix}$

$$\begin{vmatrix} 1 & 3 & 4 \\ 13 & 12 & 10 \\ 27 & 27 & 24 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 4 \\ 13 & 12 & 10 \\ 1 & 3 & 4 \end{vmatrix} \qquad r_3' = r_3 - 2r_2$$

$= 0$ (two rows equal).

(b) To evaluate $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \end{vmatrix}$

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 5 & 35 \\ 1 & 7 & 49 \end{vmatrix} \qquad r_2' = r_2 - r_1$$

$$= \begin{vmatrix} 1 & 1 & 1 \\ 0 & 5 & 35 \\ 0 & 6 & 48 \end{vmatrix} \qquad r_3' = r_3 - r_1$$

$= 1 \begin{vmatrix} 5 & 35 \\ 6 & 48 \end{vmatrix}$ (expanding by row 1,† the other minors being zero)

$= 5 \times 48 - 6 \times 35$
$= 240 - 210$
$= 30$

† Some readers will already be able to expand this by the 1st *column*.

EXERCISE **14c**

(1) Given $\Delta = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 3 & 4 \\ 4 & 6 & 8 \end{vmatrix}$ write modified determinants $\Delta_1, \Delta_2, \Delta_3$ in the following ways:

(*i*) Δ_1 has a new third row $r_3' = r_3 - 4r_1$ where these rows refer to rows of Δ.

(*ii*) Δ_2 has a new third row $r_3' = \frac{1}{2}r_3$.

(*iii*) Δ_3 has a new third row $r_3' = r_3 - 2r_2$.

Evaluate Δ, Δ_1, Δ_2 and Δ_3, expanding by any suitable row.

(2) Evaluate again the five determinants of exercise 14a, question 2 (left-hand side), using in each case appropriate row operations to induce a pair of zeros into a row before expanding by that row.

14.4 *Column operations*

Examination of the expanded form of a 3×3 determinant shows that we could just as easily have worked on a chosen column rather than a row, e.g.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}.$$

This is called expansion 'by the first column'. Minors and cofactors are obtained in exactly the same way, and, in fact, a careful look will show that the determinant would still have the same *value* if it were re-written with rows as columns, i.e. as $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$, but note that it would have been the determinant of a different matrix, known as the *transpose* of the original. We seldom wish to transpose determinants, but the possibility has an important result, viz. that *all theorems concerning rows of a determinant are also true of columns.*

EXERCISE **14d**

(1) Given $\Delta = \begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}$, show by column operations or otherwise that $x-y$, $y-z$, $z-x$ are factors.

By substitution of (unequal) values for x, y, z determine the value of k such that $\Delta = k(x-y)(y-z)(z-x)$.

(2) Show by a column operation that $a+b+c$ is a factor of $\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$.

By extracting the factor and evaluating the resulting and the original determinants show that:

$$a^3+b^3+c^3-3abc = (a+b+c)(a^2+b^2+c^2-bc-ca-ab).$$

(3) Simplify using column operations and then evaluate the determinants

$$\begin{vmatrix} 2 & 4 & 18 \\ -3 & 10 & -27 \\ 1 & 3 & 11 \end{vmatrix} \qquad \begin{vmatrix} 2 & 4 & 18 \\ -3 & 10 & -20 \\ 1 & 3 & 11 \end{vmatrix}.$$

(*Note*: If column operations *only* are used, one usually plans to get zeros into the same *row*, and thus to expand subsequently by a row.)

(4) Simplify, using any combinations of operations, and then evaluate:

$$\begin{vmatrix} 5 & 4 & 25 \\ -2 & 3 & -4 \\ 17 & 12 & 76 \end{vmatrix} \qquad \begin{vmatrix} 41 & 9 & 17 \\ 8 & 2 & 3 \\ -1 & 4 & -2 \end{vmatrix}.$$

(*Hint*: If the figures in any row or column are very large, the first step may be designed rather to reduce them to manageable size than to produce a zero immediately.)

(5) Show that $(x-y)$ is a factor of the determinant $\Delta = \begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix}$,

and hence deduce that $\Delta = k(y-z)(z-x)(x-y)(x+y+z)$. By considering the coefficient of yz^3 in both expressions determine the value of k.

(6) Evaluate the determinants:

$$(a)\ \begin{vmatrix} 3 & 1 & -1 \\ 4 & 2 & 6 \\ 1 & 2 & 0 \end{vmatrix} \qquad (b)\ \begin{vmatrix} 8 & 2 & -4 \\ -3 & 0 & 2 \\ 4 & 1 & -6 \end{vmatrix} \qquad (c)\ \begin{vmatrix} 6 & 2 & -1 \\ 4 & 3 & -2 \\ -6 & 1 & 4 \end{vmatrix}$$

$$(d)\ \begin{vmatrix} -2 & -3 & 1 \\ -2 & -4 & 1 \\ 6 & 8 & 4 \end{vmatrix} \qquad (e)\ \begin{vmatrix} 18 & 16 & 2 \\ 7 & 5 & 1 \\ 2 & 1 & -1 \end{vmatrix} \qquad (f)\ \begin{vmatrix} 17 & 19 & 18 \\ 5 & 6 & 6 \\ 2 & 2 & 2 \end{vmatrix}$$

14.5 *The determinant 'determining' the nature of equations*

When we attempt to solve a set of simultaneous linear equations such as

$$\begin{aligned} 3x+4y+\ \ z &= 8 \\ 2x+2y-\ 3z &= 0 \\ 3x+2y-10z &= 7 \end{aligned}$$

and we write the equations in detached coefficient form, i.e.

$$\left[\begin{array}{rrr|r} 3 & 4 & 1 & 8 \\ 2 & 2 & -3 & 0 \\ 3 & 2 & -10 & 7 \end{array}\right]$$

for ease of manipulation, all the operations that we might wish to perform have been discussed using determinants. As an illustration, we set about the equations above in order to solve them, letting Δ be the value of $\begin{vmatrix} 3 & 4 & 1 \\ 2 & 2 & -3 \\ 3 & 2 & -10 \end{vmatrix}$ and consider what happens to Δ as we make each step. The work is set out fully as follows:

3	4	1	8	
2	2	−3	0	
3	2	−10	7	
3	4	1	8	
2	2	−3	0	Value of determinant $= \Delta$
0	−2	−11	−1	$r_3' = r_3 - r_1$
6	8	2	16	$r_1' = r_1 \times 2$
6	6	−9	0	$r_2' = r_2 \times 3$
0	−2	−11	−1	New value of determinant $= 6\Delta$
6	8	2	16	
0	−2	−11	−16	$r_2' = r_2 - r_1$
0	−2	−11	−1	Value of determinant unchanged $= 6\Delta$
6	8	2	16	
0	−2	−11	−16	
0	0	0	15	$r_3' = r_3 - r_2$
				Value of determinant unchanged $= 6\Delta$

But we now see that the value of the determinant of the existing equations is zero. Hence, the value of the original determinant is zero.

In fact, each of the processes of elimination in the equations

$$\begin{aligned} a_1x+b_1y+c_1z &= d_1 \\ a_2x+b_2y+c_2z &= d_2 \qquad (\mathbf{d} \neq 0) \\ a_3x+b_3y+c_3z &= d_3 \end{aligned}$$

only involves multiplying Δ by a certain factor which is non-zero, so that $\Delta = 0$ before elimination $\Leftrightarrow$ $\Delta = 0$ after elimination. Hence, the vanishing of Δ as defined at the beginning of the chapter is the condition that we require for a set of equations to be regular. (A different approach to the same situation is given in the next section.)

14.6 A set of linear equations can be considered as a vector equation, e.g. the triad discussed in

$$x\begin{pmatrix}3\\2\\1\end{pmatrix}+y\begin{pmatrix}4\\2\\2\end{pmatrix}+z\begin{pmatrix}1\\-3\\4\end{pmatrix}=\begin{pmatrix}8\\0\\7\end{pmatrix}$$

and the reader will remember that the condition for a unique solution was then that the vectors

$$\begin{pmatrix}3\\2\\1\end{pmatrix},\quad \begin{pmatrix}4\\2\\2\end{pmatrix}\quad \text{and} \quad \begin{pmatrix}1\\-3\\4\end{pmatrix}$$

should be linearly independent, i.e. that there should be *no* numbers λ, μ and ν (other than all-zero) such that

$$\lambda\begin{pmatrix}3\\2\\1\end{pmatrix}+\mu\begin{pmatrix}4\\2\\2\end{pmatrix}+\nu\begin{pmatrix}1\\-3\\4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

If values λ, μ and ν *do* exist, and we let Δ be the value of

$$\begin{vmatrix}3 & 4 & 1\\2 & 2 & -3\\1 & 2 & 4\end{vmatrix},$$

$$\begin{aligned} \text{then} \quad \lambda\Delta &= \begin{vmatrix}3\lambda & 4 & 1\\2\lambda & 2 & -3\\\lambda & 2 & 4\end{vmatrix} \\ &= \begin{vmatrix}3\lambda+4\mu+\nu & 4 & 1\\2\lambda+2\mu-3\nu & 2 & -3\\\lambda+2\mu+4\nu & 2 & 4\end{vmatrix} \quad \text{using column properties} \\ &= \begin{vmatrix}0 & 4 & 1\\0 & 2 & -3\\0 & 2 & 4\end{vmatrix} \quad \text{for we have supposed that } \lambda, \mu, \nu \text{ exist to make the vectors linearly dependent} \\ &= 0. \end{aligned}$$

Hence if a set of three 3-vectors is linearly dependent, then the determinant of those vectors is zero. Conversely, if $\Delta \neq 0$, it is *not* possible to make one column vanish by using column operations involving multiples of columns, and there are no numbers λ, μ, ν such that the vectors are linearly dependent. Hence they are independent.

We see that $\Delta = 0 \Leftrightarrow$ three 3-vectors linearly dependent.

Note: Here we have proved that $\Delta = 0 \Leftrightarrow$ three column 3-vectors are linearly dependent, but the structure of Δ implies that the row-vectors are also linearly dependent. Hence not only are

$\begin{pmatrix}3\\2\\1\end{pmatrix}, \begin{pmatrix}4\\2\\2\end{pmatrix}$ and $\begin{pmatrix}1\\-3\\4\end{pmatrix}$ linearly dependent, but so are $\begin{pmatrix}3\\4\\1\end{pmatrix}, \begin{pmatrix}2\\2\\-3\end{pmatrix}$ and $\begin{pmatrix}1\\2\\4\end{pmatrix}$

derived from the row-vectors $(3, 4, 1)$, $(2, 2, -3)$ and $(1, 2, 4)$. This means that the l.h.s. of one equation will be a composite version of the l.h.s. of the other two. The implications of this are discussed in chapter 17.

14.7 *Geometrical properties of determinants*

At this point it is worth while making a few comments about the geometric properties of the determinant as applied to linear transformations. The 2×2 case illustrates this particularly well.

The matrix $\mathbf{M} = \begin{pmatrix}a & b\\c & d\end{pmatrix}$ is an operator representing a linear mapping which transforms the unit square into a parallelogram as shown in the diagram.

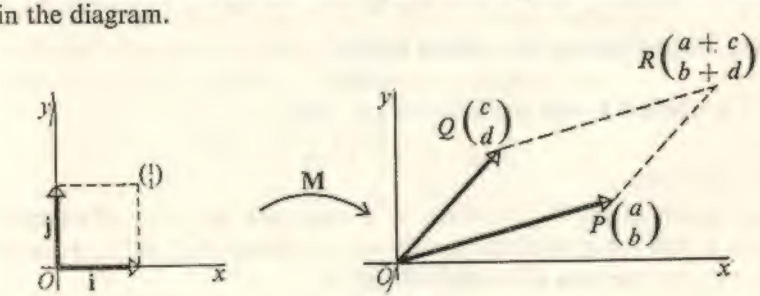


FIG. 58

The area of the parallelogram may be seen by completing the figure as shown at the top of the next page.

$$\begin{aligned} \text{Area} \quad OPRQ &= (a+c)(b+d)-2bc-cd-ab \\ &= ab+bc+ad+cd-2bc-cd-ab \\ &= ad-bc \\ &= \det \mathbf{M}. \end{aligned}$$

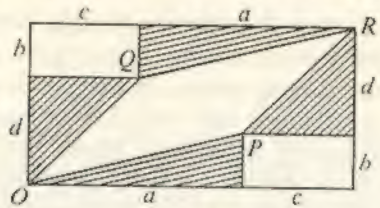


FIG. 59

(*Note*: A particularly convenient diagram has been drawn to achieve this result. The reader should try various cases to decide on the significance of det **M** being negative, as of course it might be.)

Similarly, it may be shown, but not easily (unless the reader is familiar with scalar triple products of vectors) that the volume of the parallelepiped into which the unit cube is mapped by the matrix $\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ is also det **M**.

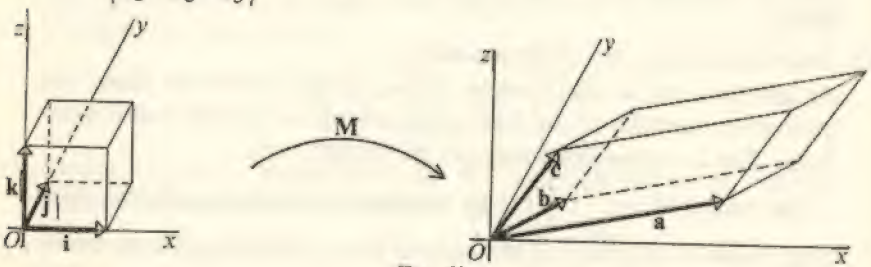


FIG. 60

Note: In the diagram the column vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ is labelled **a**; and similarly for $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ and $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$.

More geometry

If we return to the 2×2 matrix, it is clear that the area of triangle OP_1P_2 is half the area of the parallelogram. Hence if P_1 is (x_1, y_1) and P_2 is (x_2, y_2) the area of triangle OP_1P_2 is

$$\tfrac{1}{2} \times \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

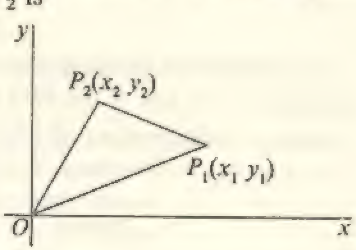


FIG. 61

(*Note*: Unless the sense of travel from O to P_1 to P_2 to O is anti-clockwise, the area will be negative.)

We extend this to general case—see Fig. 62—in the obvious way by saying $\Delta P_1P_2P_3 = \Delta OP_1P_2 - \Delta OP_1P_3 - \Delta OP_3P_2$ (or the last term could be $+\Delta OP_2P_3$).

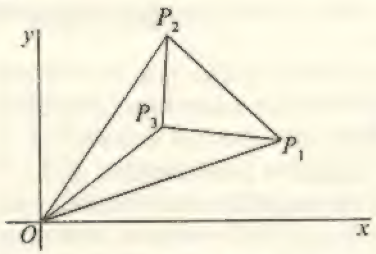


FIG. 62

Thus, $\Delta P_1P_2P_3 = \tfrac{1}{2}\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} - \tfrac{1}{2}\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + \tfrac{1}{2}\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}$.

This is just $\tfrac{1}{2}\times$expansion of a 3×3 determinant by the l.h. column where all the elements in that column are 1.

Therefore

$$\Delta P_1P_2P_3 = \tfrac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} \text{ or, more commonly, } \tfrac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

the result being numerically correct in all cases, with a sign which is positive if the sense P_1 to P_2 to P_3 to P_1 is anti-clockwise.

If the area of the triangle formed is zero, we get the condition that (x_1, y_1), (x_2, y_2) and (x_3, y_3) should be collinear, viz.

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

EXERCISE 14c

(1) (*a*) Prove that the points (6, 1) (10, 2) and $(-2, -1)$ are collinear.

(*b*) Find the area of the triangle ABC formed by the points $(-2, 5)$, (7, 6), (5, 10). Does the sign of your determinant indicate whether the circuit $A \to B \to C \to A$ is anti-clockwise? (*Hint*: It is possible to perform two simple row operations such as $r_2' = r_2 - r_1$ and $r_3' = r_3 - r_1$ in one move, but this is only safe if one such move does not affect the other.)

(2) What is the volume of the parallelepiped which has the origin at one corner, and has as adjacent sides the lines joining the origin to (1, 2, 3), (2, 3, 7) and (7, 5, 1)?

(3) Show directly and by determinant that three column vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ of order 3 will always be linearly dependent if one of them (e.g. $\mathbf{w}$) has all its components zero. Give a geometric interpretation of the space spanned by $\mathbf{u}$ and $\mathbf{v}$.

(4) Show that four column vectors $\mathbf{t}$, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ of order 3 will always be linearly dependent. (*Hint*: Consider separately the cases in which $\mathbf{t}$, $\mathbf{u}$, $\mathbf{v}$ are linearly dependent and those in which they are not, and try to solve $x\mathbf{t}+y\mathbf{u}+z\mathbf{v} = \mathbf{w}$.) Suggest a generalisation of this result.

(5) By considering the magnification produced when transforming the simple unit square, first by 2×2 matrix $\mathbf{A}$, and secondly by the product matrix $\mathbf{BA}$, and using the result $\det(\mathbf{BA}) = \det\mathbf{B}\times\det\mathbf{A}$, show that $\det\mathbf{B}$ is the area-magnifying factor involved in transforming *any* parallelogram with a vertex at O (i.e. not only a unit square).

(6) Show that if O is the origin and P, Q are points of an integer-lattice (i.e. that both have integral coefficients) then when the parallelogram $OPRQ$ is completed,

(*i*) the point R has integral coefficients, and

(*ii*) the parallelogram has integral area.

(7) If in question 6 the point P is (m, n), Q is (p, q) and $mq-np = +1$, show that there is no other lattice point inside $OPRQ$.

(8) Show that if $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are 3-vectors forming the columns of a matrix $\mathbf{M}$, and $\mathbf{w} = \lambda\mathbf{u}+\mu\mathbf{v}$, then $\det\mathbf{M} = 0$.

chapter 15

'ELEMENTARY' MATRICES

15.1 So far we have considered matrices only as machines for performing linear mappings or as arrays of numbers representing the coefficients in an equation. We now begin to look at matrices in themselves, in order to find out more about their character and habits. It is quite usual for the pure mathematician when he studies any particular concept to try to classify the individual elements, in this case, matrices, into types, and we follow this for the moment.

But, to begin, we need some tools, and these tools are called *elementary* matrices. They will not be defined formally until the reader has gained some first-hand experience of them. (*Note*: The word 'elementary' is here being used in a technical sense, not as a synonym for 'simple'; it means simple in a special way which will be shown by the following exercise for the reader.)

EXERCISE 15a

In all these questions $\mathbf{A}$ is the matrix $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$.

(1) Consider the product $\mathbf{EA}$ for the various matrices $\mathbf{E}$ which are given, and state in words the effect of having multiplied by $\mathbf{E}$.

(*a*) $\mathbf{E} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*b*) $\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(*c*) $\mathbf{E} = \begin{pmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*d*) $\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(*e*) $\mathbf{E} = \begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*f*) $\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(2) By what matrices $\mathbf{E}_1$ and $\mathbf{E}_2$ would you pre-multiply $\mathbf{A}$ if you wished to:

(*a*) interchange row 1 with row 3?
(*b*) add $k \times$(row 1) to row 3?

Notice that your answers should in each case be the same as if the stated operation had been applied to $\mathbf{I}$.

(3) Form the products $\mathbf{E}_1\mathbf{E}_2$, and $\mathbf{E}_2\mathbf{E}_1$, where $\mathbf{E}_1$ and $\mathbf{E}_2$ are the matrices from question 2. State in words what effect you think would be obtained by multiplying $\mathbf{E}_1\mathbf{E}_2$ and $\mathbf{E}_2\mathbf{E}_1$ into $\mathbf{A}$, and then compute $\mathbf{E}_1\mathbf{E}_2\mathbf{A}$ and $\mathbf{E}_2\mathbf{E}_1\mathbf{A}$ to verify that what you suggested is right.

(4) Give 2×2 matrices which have similar row properties to the properties possessed by these 3×3 matrices.

We are now in a position to define an elementary matrix. An *elementary matrix* is a matrix which, by pre-multiplying any other matrix,† performs a row operation on that matrix.

It is clearly important to be able to associate each type of row operation with its appropriate matrix. This can be done by means of a simple observation, which the reader should verify from his worked examples in exercise 15a, viz. the matrix required is *in every case the result of applying the row operation to the unit matrix* $\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$ as follows.

The reader should verify in each case, by forming the product $\mathbf{EA}$, that the elementary matrix performs the required operation.

Type (a): Interchange of two rows

e.g. of first and second rows, i.e. $r_1' = r_2, r_2' = r_1$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \longrightarrow \begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix},$$

and in verification

$$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{pmatrix} = \begin{pmatrix}a_2&b_2&c_2\\a_1&b_1&c_1\\a_3&b_3&c_3\end{pmatrix},$$

i.e. multiplication by $\mathbf{E}$ carries out the interchange.

† Provided, of course, that they are conformable for this multiplication.

Type (b): To add a multiple of one row to another

e.g. $r_1' = r_1 + \lambda r_2$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \longrightarrow \begin{pmatrix}1&\lambda&0\\0&1&0\\0&0&1\end{pmatrix}.$$

Verify the character of this matrix by forming $\mathbf{EA}$.

Type (c): To multiply a row by a factor μ

e.g. $r_3' = \mu r_3$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \longrightarrow \begin{pmatrix}1&0&0\\0&1&0\\0&0&\mu\end{pmatrix}.$$

Again verify the character of this elementary matrix.

15.2 *Echelon form*

We have seen earlier (in chapter 5) that any 3×3 matrix $\mathbf{A}$ can be put, by row operations, into what we called an echelon form, in which all elements below the main diagonal are zero.

It will be remembered that the connection between the echelon matrix and its original was their relevance to the same equation-solving situation. The reader may see at a later stage another way of expressing this equivalence: it is sufficient here that the new matrix can be obtained from the old by multiplying by successive elementary matrices.

In this section we shall carry out the operations successively on $\mathbf{A}$ in the same way as we did previously, but alongside for information we shall also carry out the same operations successively on $\mathbf{I}$, thereby showing at each stage the correct product of all the elementary matrices used; e.g. after 3 operations we get $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{I}$ which is $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$. The reduced matrix at this stage will of course be $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$. (Notice that after $\mathbf{E}_1$ no other elementary matrix will appear in the work, but only the cumulative product of them.)

Worked example: To reduce the matrix $\begin{pmatrix}3&4&2\\1&6&1\\3&4&-2\end{pmatrix}$ to an echelon form, showing at each stage the total multiplying matrix used.

$$\text{Initially: } \mathbf{A} = \begin{pmatrix}3&4&2\\1&6&1\\3&4&-2\end{pmatrix} \qquad \mathbf{I} = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

Operation	*Result*	*Result*	
Interchange rows 1 and 2 ($r_1' = r_2$, $r_2' = r_1$)	$\begin{pmatrix} 1 & 6 & 1 \\ 3 & 4 & 2 \\ 3 & 4 & -2 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	
$r_3' = r_3 - r_2$	$\begin{pmatrix} 1 & 6 & 1 \\ 3 & 4 & 2 \\ 0 & 0 & -4 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$	(Note that this is $\mathbf{E}_1\mathbf{E}_2$ and is not itself an elementary matrix.)
$r_2' = r_2 - 3r_1$	$\mathbf{B} = \begin{pmatrix} 1 & 6 & 1 \\ 0 & -14 & -1 \\ 0 & 0 & -4 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$	

This final matrix **B** is now said to be 'in echelon form', the characteristic being that *there are only zeros underneath the main diagonal.* If the original matrix had been rectangular, say,

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

then the echelon form would be

$$\begin{pmatrix} . & . & . & . \\ 0 & . & . & . \\ 0 & 0 & . & . \end{pmatrix}$$

where the dots may or may not be zero. For a 4×3 matrix the echelon form would be

$$\begin{pmatrix} . & . & . \\ 0 & . & . \\ 0 & 0 & . \\ 0 & 0 & 0 \end{pmatrix}$$

where again the dots may or may not be zero.

(*Note*: The original reduction might have been carried out in a different way, just as in solving linear equations the elimination may be done in a different order. Thus there is no unique 'correct' echelon form; but certain features will be found in common, notably that if one row is all zeros for one form, this will be so for all.)

15.3 We can formalise the processes of the previous section if we call the elementary matrices used $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_3$, and the final reduced matrix **B**; then $$\mathbf{B} = \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

The important thing here is the final form, not the particular matrices used; provided every step is a row operation the final matrix will be given by an expression of the form:

Final matrix = (product of elementary matrices) $\times$ (initial matrix)

It is not necessary therefore to write down the individual elementary matrices; it will suffice to perform the operations. It is, however, easy to make mistakes, and it is certainly worth noting what operations have been performed so that the work can be easily checked.

Example (*a*):

Reduce $\begin{pmatrix} 3 & 4 & 2 \\ 1 & 6 & 1 \\ 3 & 4 & -2 \end{pmatrix}$ to echelon form.

It will save fractions to change row 1 with row 2 first. Accordingly:

$$\begin{pmatrix} 3 & 4 & 2 \\ 1 & 6 & 1 \\ 3 & 4 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 6 & 1 \\ 3 & 4 & 2 \\ 3 & 4 & -2 \end{pmatrix} \quad \begin{matrix} r_1' = r_2 \\ r_2' = r_1 \end{matrix}$$

where the $\longrightarrow$ indicates that the second matrix is derived from the first.

Then $$\begin{pmatrix} 1 & 6 & 1 \\ 3 & 4 & 2 \\ 3 & 4 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 6 & 1 \\ 3 & 4 & 2 \\ 0 & 0 & -4 \end{pmatrix} \quad r_3' = r_3 - r_2$$

$$\longrightarrow \begin{pmatrix} 1 & 6 & 1 \\ 0 & -14 & -1 \\ 0 & 0 & -4 \end{pmatrix} \quad r_2' = r_2 - 3r_1.$$

Whether in connection with linear equations or not, it is quite common practice to convert a *non*-square matrix (or even a collection of vectors) into an echelon form. Our next example will show this; it will be seen that several elementary operations can be combined in one move.

Example (*b*):

To reduce $\begin{pmatrix} 1 & 2 & 1 \\ 4 & 7 & -1 \\ 2 & 1 & 3 \\ 1 & 4 & 2 \end{pmatrix}$ to echelon form.

Combining operations, we have:

$$\begin{pmatrix} 1 & 2 & 1 \\ 4 & 7 & -1 \\ 2 & 1 & 3 \\ 1 & 4 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 0 & -3 & 1 \\ 0 & 2 & 1 \end{pmatrix} \quad \begin{matrix} \\ r_2' = r_2 - 4r_1 \\ r_3' = r_3 - 2r_1 \\ r_4' = r_4 - r_1 \end{matrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 0 & 0 & 16 \\ 0 & 0 & -9 \end{pmatrix} \quad \begin{matrix} \\ \\ r_3' = r_3 - 3r_2 \\ r_4' = r_4 + 2r_2 \end{matrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 0 & 0 & 16 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{matrix} \\ \\ \\ r_4' = r_4 + (\tfrac{9}{16})r_3. \end{matrix}$$

EXERCISE 15b

(1) Reduce the following matrices to echelon form:

$(a) \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$, $(b) \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix}$, $(c) \begin{pmatrix} 0 & 5 \\ 4 & 10 \end{pmatrix}$, $(d) \begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix}$.

(2) Reduce $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to echelon form.

Is it possible for the reduced form to have zero elements only, in the second row? If so, give the condition for this to be so in terms of a, b, c and d. Is your working valid when $a = 0$? If not, put $a = 0$ and rework the question. What is your final conclusion?

(3) Reduce these matrices to echelon form:

$(a) \begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 1 \\ 3 & 1 & 1 \end{pmatrix}$ $(b) \begin{pmatrix} 2 & 0 & 3 \\ 1 & 4 & 0 \\ 4 & 8 & 3 \end{pmatrix}$ $(c) \begin{pmatrix} 4 & 2 & 3 \\ -1 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$

$(d) \begin{pmatrix} 0 & 1 & 2 \\ 0 & 4 & 6 \\ 1 & 1 & 1 \end{pmatrix}$ $(e) \begin{pmatrix} 3 & 1 & 1 & 2 \\ 2 & 4 & 1 & 3 \\ 1 & 7 & 1 & 4 \end{pmatrix}$ $(f) \begin{pmatrix} 1 & 2 & 6 & -4 \\ 3 & 1 & 0 & 2 \\ 5 & 0 & -6 & 0 \\ 0 & 5 & 18 & -4 \end{pmatrix}$

(4) It is possible to avoid fractions when reducing a matrix with integer elements to echelon form by multiplying a row before subtracting a multiple of another row. For example, in reducing $\begin{pmatrix} 7 & 2 \\ -5 & -3 \end{pmatrix}$ to echelon form we might say

$$\begin{pmatrix} 7 & 2 \\ -5 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 7 & 2 \\ 0 & -11 \end{pmatrix} \quad r_2' = 7r_2 + 5r_1.$$

Justify this in terms of elementary matrices and then reduce

$\begin{pmatrix} 3 & 7 & 2 \\ 4 & 6 & -3 \\ 5 & 5 & -8 \end{pmatrix}$ to echelon form.

15.4 *Singular and non-singular matrices*

So far the elementary matrices have split all possible matrices into two classes, according to whether the echelon form has a final row of zeros or not. The reader will no doubt have realised that the condition for zeros in the final row for square matrices is that the determinant of the matrix is zero, i.e. that the matrix is *singular*. If the determinant of a square matrix is not zero, i.e. the final row of the echelon form of a square matrix does not consist of zeros, then the matrix is *non-singular*.

We now restrict our attention to the non-singular matrices, pushing the use of elementary matrices further. As in the reduction to echelon form, we proceed with an example.

We consider the matrix **A**, reduce it to echelon form, and then further, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 12 \\ 3 & 0 & 10 \end{pmatrix}.$$

Then

$$\mathbf{A} \longrightarrow \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 2 \\ 0 & -3 & -5 \end{pmatrix} \quad \begin{matrix} r_2' = r_2 - 2r_1 \\ r_3' = r_3 - 3r_1 \end{matrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad r_3' = r_3 + 3r_2.$$

We now continue so that

$$\begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{matrix} r_1' = r_1 - 5r_3 \\ r_2' = r_2 - 2r_3 \end{matrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad r_1' = r_1 - r_2$$

and the original matrix $\mathbf{A}$ has been reduced, using row operations only, to the unit matrix, $\mathbf{I}$. But each row operation is equivalent to pre-multiplication by an elementary matrix, so what we have done is to say $(\mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_2\mathbf{E}_1)\mathbf{A} = \mathbf{I}$.

But this is of the form $\mathbf{BA} = \mathbf{I}$ where

$$\mathbf{B} = \mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_2\mathbf{E}_1$$
$$\text{or} \quad \mathbf{B} = (\mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_2\mathbf{E}_1)\mathbf{I}.$$

Multiplying $\mathbf{I}$ by this succession of matrices involves the same row operations on $\mathbf{I}$ as on $\mathbf{A}$. In this case—the reader is advised to check—these operations to $\mathbf{I}$ give:

$$\mathbf{B} = \begin{pmatrix} 30 & -10 & -3 \\ 16 & -5 & -2 \\ -9 & 3 & 1 \end{pmatrix}$$

and direct computation gives:

$$\mathbf{BA} = \begin{pmatrix} 30 & -10 & -3 \\ 16 & -5 & -2 \\ -9 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 12 \\ 3 & 0 & 10 \end{pmatrix}$$

$$= \begin{pmatrix} 30-20-9, & 30-30, & 150-120-30 \\ 16-10-6, & 16-15, & 80-60-20 \\ -9+6+3, & -9+9, & -45+36+10 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}.$$

This matrix $\mathbf{B}$ is the inverse matrix which was discussed earlier. It is clear that if we started from $\mathbf{B}$ and used elementary row operations to form $\mathbf{I}$, then performing the same operations on $\mathbf{I}$ would lead to $\mathbf{A}$.

Hence $\mathbf{BA} = \mathbf{I} \Leftrightarrow \mathbf{AB} = \mathbf{I}$.

We can also prove that $\mathbf{B}$ is unique. For suppose that two matrices $\mathbf{B}$ and $\mathbf{C}$ have the property that $\mathbf{BA} = \mathbf{AB} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{AC} = \mathbf{I}$.

Then

$$\begin{aligned} \mathbf{B} &= \mathbf{BI} \\ &= \mathbf{B}(\mathbf{AC}) \\ &= (\mathbf{BA})\mathbf{C} \\ &= \mathbf{IC} \\ &= \mathbf{C} \end{aligned}$$

and the result is proved. The notation used for the inverse matrix is the usual one, namely $\mathbf{B} = \mathbf{A}^{-1}$.

The method employed to find the inverse matrix can be improved by better organisation. For, remembering that the same operations are to be performed on the original matrix and on $\mathbf{I}$, it is better to have the two side by side, and to deal with the two together.

Example

Find the inverse of $\quad \mathbf{A} = \begin{pmatrix} 1 & -1 & -3 \\ -1 & 5 & 14 \\ 2 & -4 & -11 \end{pmatrix}$

$$\begin{pmatrix} 1 & -1 & -3 \\ -1 & 5 & 14 \\ 2 & -4 & -11 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 3 \\ 0 & 4 & 11 \\ 0 & -2 & -5 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \qquad \begin{matrix} r_2' = r_2 + r_1 \\ r_3' = r_2 - 2r_1 \end{matrix}$$

$$\begin{pmatrix} 1 & -1 & 3 \\ 0 & 4 & 11 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -3 & 1 & 2 \end{pmatrix} \qquad r_3' = 2r_3 + r_2$$

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} -8 & 3 & 6 \\ 34 & -10 & -22 \\ -3 & 1 & 2 \end{pmatrix} \qquad \begin{matrix} r_1' = r_1 - 3r_3 \\ r_2' = r_2 - 11r_3 \end{matrix}$$

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 2 & 2 & 2 \\ 34 & -10 & -22 \\ -3 & 1 & 2 \end{pmatrix} \qquad r_1' = 4r_1 + r_2$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{17}{2} & -\frac{5}{2} & -\frac{11}{2} \\ -3 & 1 & 2 \end{pmatrix} \qquad \begin{matrix} r_1' = r_1/4 \\ r_2' = r_2/4 \end{matrix}$$

Hence $\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{17}{2} & -\frac{5}{2} & -\frac{11}{2} \\ -3 & 1 & 2 \end{pmatrix}$.

As a check—always advisable:

$$\mathbf{A}^{-1}\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{17}{2} & -\frac{5}{2} & -\frac{11}{2} \\ -3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 & -3 \\ -1 & 5 & 14 \\ 2 & -4 & -11 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}$$

When manipulating matrices in this way, it is nearly always worth while avoiding fractions until the last possible point, by using suitable multipliers.

It is possible to shorten the process still further by not making the original matrix pass through the echelon stage. We give an example:

Find the inverse of $\mathbf{A} = \begin{pmatrix} -1 & -2 & 1 \\ 7 & 1 & 1 \\ -2 & -2 & 1 \end{pmatrix}$

$$\begin{pmatrix} -1 & -2 & 1 \\ 7 & 1 & 1 \\ -2 & -2 & 1 \end{pmatrix} \begin{array}{c} | \\ | \\ | \end{array} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & -2 & 1 \\ 0 & -13 & 8 \\ 0 & 2 & -1 \end{pmatrix} \begin{array}{c} | \\ | \\ | \end{array} \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} \\ r_2' = r_2+7r_1 \\ r_3' = r_3-2r_1 \end{array}$$

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 2 & -1 \end{pmatrix} \begin{array}{c} | \\ | \\ | \end{array} \begin{pmatrix} -1 & 0 & 1 \\ -9 & 1 & 8 \\ -2 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} r_1' = r_1+r_3 \\ r_2' = r_2+8r_3 \\ \end{array}$$

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{array}{c} | \\ | \\ | \end{array} \begin{pmatrix} -1 & 0 & 1 \\ -9 & 1 & 8 \\ 12 & -2 & -13 \end{pmatrix} \quad \begin{array}{l} \\ \\ r_3' = 3r_3-2r_2 \end{array}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{array}{c} | \\ | \\ | \end{array} \begin{pmatrix} 1 & 0 & -1 \\ -3 & \frac{1}{3} & \frac{8}{3} \\ -4 & \frac{2}{3} & \frac{13}{3} \end{pmatrix} \quad \begin{array}{l} r_1' = -r_1 \\ r_2' = r_2/3 \\ r_3' = -r_3/3 \end{array}$$

Hence $\mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -3 & \frac{1}{3} & \frac{8}{3} \\ -4 & \frac{2}{3} & \frac{13}{3} \end{pmatrix}$

and checking:

$$\mathbf{A}^{-1}\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ -3 & \frac{1}{3} & \frac{8}{3} \\ -4 & \frac{2}{3} & \frac{13}{3} \end{pmatrix}\begin{pmatrix} -1 & -2 & 1 \\ 7 & 1 & 1 \\ -2 & -2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1+2 & -2+2 & 1-1 \\ 3+\frac{7}{3}-\frac{16}{3} & 6+\frac{1}{3}-\frac{16}{3} & -3+\frac{1}{3}+\frac{8}{3} \\ 4+\frac{14}{3}-\frac{26}{3} & 8+\frac{2}{3}-\frac{26}{3} & -4+\frac{2}{3}+\frac{13}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}.$$

This is one of the most efficient ways of finding the inverse of a matrix. It is not usually the best way for a small matrix—we shall see that in the next chapter—but it is easy to programme, and this is its greatest value.

Notice that, when inverting, if the echelon stage—either a lower or an upper echelon—is reached, one of the diagonal elements is zero, there will be *no* inverse matrix. Look out for this in the following examples, and check for yourself why this should be so.

EXERCISE 15c

(1) Find the inverse of each of the following matrices:

(*i*) $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ (*ii*) $\begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix}$ (*iii*) $\begin{pmatrix} 4 & 3 \\ 9 & 7 \end{pmatrix}$

Can you see a simple rule appearing? If you can formulate it, do so.

(2) If you managed to formulate a rule in question 1, try it for inverting the following matrices. If you did not, use the method of elementary row operations. After using your rule, check your answer; if anything goes wrong, invert in the usual way.

(*i*) $\begin{pmatrix} 4 & -9 \\ 1 & -2 \end{pmatrix}$ (*ii*) $\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$ (*iii*) $\begin{pmatrix} 5 & 2 \\ 11 & 3 \end{pmatrix}$

(*iv*) $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$ (*v*) $\begin{pmatrix} 3 & 1 \\ 7 & 1 \end{pmatrix}$ (*iv*) $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

(3) Make all the matrices in questions 1 and 2 fall into a pattern by finding by elementary row operations the inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(4) Invert, where possible

(*i*) $\begin{pmatrix} 1 & -2 & 1 \\ 2 & -3 & 0 \\ -1 & 5 & -6 \end{pmatrix}$ (*ii*) $\begin{pmatrix} 3 & -4 & 2 \\ 1 & 5 & -2 \\ 1 & -2 & 1 \end{pmatrix}$

(*iii*) $\begin{pmatrix} 3 & 1 & -4 \\ 5 & 0 & -1 \\ -1 & -2 & 7 \end{pmatrix}$ (*iv*) $\begin{pmatrix} 2 & -5 & 9 \\ 1 & -2 & 2 \\ 1 & -3 & 5 \end{pmatrix}$.

(5) Show that the following elementary matrices—one of each type—are non-singular, and that the inverse of each is another elementary matrix:

(*i*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ (*ii*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*iii*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}$

where $\mu \neq 0$.

15.5 *Non-integral coefficients*

The method of row operations is the one used in practice to invert matrices with non-integral coefficients. The layout is very similar to the layout already used for integral coefficients, except that, to guard against blunders, we use an additional column as a sum check (see section 4.5). We first of all arrange, before subtracting a multiple of a row from another row, that the largest coefficient in the first column is in the leading position, corresponding to the pivotal equation discussed in section 4.5.

As an example, we invert the matrix $\begin{pmatrix} 0{\cdot}5000 & 0{\cdot}3200 \\ 0{\cdot}1000 & 0{\cdot}7000 \end{pmatrix}$.

Original matrix				*Sum check*		
0·5000	0·3200	1·0000	0·0000	1·8200		
0·1000	0·7000	0·0000	1·0000	1·8000		
0·5000	0·3200	1·0000	0·0000	1·8200	√	$r_1' = r_1$
0·0000	0·6360	−0·2000	1·0000	1·4360	√	$r_2' = r_2 - 0{\cdot}2000\, r_1$
0·5000	0·0001	1·1006	−0·5031	1·0976	√	$r_1' = r_1 - 0{\cdot}5031\, r_2$
0·0000	0·6360	−0·2000	1·0000	1·4360	√	
1·0000	0·0002	2·2012	−1·0062	2·1952	√	$r_1' = 2{\cdot}0000\, r_1$
0·0000	0·9999	−0·3144	1·5721	2·2576	√	$r_2' = 1{\cdot}5721\, r_2$

Inverse matrix

The inverse matrix is seen to be $\begin{pmatrix} 2{\cdot}2012 & -1{\cdot}0062 \\ -0{\cdot}3144 & 1{\cdot}5721 \end{pmatrix}$ and direct multiplication with the original matrix gives

$$\begin{pmatrix} 0{\cdot}5000 & 0{\cdot}3200 \\ 0{\cdot}1000 & 0{\cdot}7000 \end{pmatrix} \begin{pmatrix} 2{\cdot}2012 & -1{\cdot}0062 \\ -0{\cdot}3144 & 1{\cdot}5721 \end{pmatrix} = \begin{pmatrix} 1{\cdot}0000 & 0{\cdot}0000 \\ 0{\cdot}0000 & 0{\cdot}9999 \end{pmatrix}.$$

EXERCISE 15d

Using a desk calculating machine, find the inverses of the following two matrices:

(1) $\begin{pmatrix} 0{\cdot}4000 & 0{\cdot}1000 \\ 0{\cdot}2410 & 0{\cdot}6000 \end{pmatrix}$. (2) $\begin{pmatrix} 0{\cdot}3000 & -0{\cdot}1600 \\ 0{\cdot}1250 & 0{\cdot}7000 \end{pmatrix}$.

chapter 16

BACK TO DETERMINANTS

16.1 We have already defined *cofactors* and shown how to form them from the minors of a determinant.

Considering

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

the cofactor corresponding to a_1 is called A_1 and

$$A_1 = +\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}.$$

Similarly,

$$B_1 = -\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}.$$

Using this notation, the expressions for the expansion of the determinant by any row or column were found to take a very simple appearance, e.g.

$$\begin{aligned} \Delta &= a_1A_1+b_1B_1+c_1C_1 && \text{(by the first row)} \\ &= b_1B_1+b_2B_2+b_3B_3 && \text{(by the second column).} \end{aligned}$$

EXERCISE **16a**

(1) Let Δ be the determinant

$$\begin{vmatrix} 3 & 1 & 7 \\ 2 & 3 & -1 \\ 1 & 4 & -2 \end{vmatrix}.$$

Evaluate the cofactors corresponding to each element. If the cofactors of the first column are called A_1, A_2 and A_3, evaluate:

$$1A_1+3A_2+4A_3 \quad \text{and} \quad 7A_1-A_2-2A_3.$$

If the cofactors of the first row are A_1, B_1 and C_1 evaluate:

$$2A_1+3B_1-C_1 \quad \text{and} \quad A_1+4B_1-2C_1.$$

Can you generalise these results?

(2) If Δ is the determinant

$$\begin{vmatrix} 2 & 3 & 5 \\ 2 & 6 & -8 \\ 1 & 3 & -2 \end{vmatrix}$$

find the cofactors of each element, and make up a new determinant out of these cofactors. Evaluate both Δ and this new determinant, Δ'. Is your result a coincidence?

Now evaluate the cofactors of Δ'. Can you see a pattern developing?

16.2 *Alien cofactors*

In the previous exercise the reader will have noticed a curious property of cofactors, viz. that when we take the elements of a row (or column) and multiply by its own cofactors we get Δ, but if we take the cofactors for some other row (or column)—'alien cofactors', we shall call them—we get zero.

For, if

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

then $A_1 = +\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$, $B_1 = -\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$, $C_1 = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$.

The expression $a_2A_1+b_2B_1+c_2C_1$ is then identical with the expansion

$$\begin{vmatrix} a_2 & b_2 & c_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

which is zero (two rows equal).

16.3 *Cramer's rule for linear equations*

We consider the equations:

$$\begin{aligned} a_1x+b_1y+c_1z &= d_1 && \ldots (1) \\ a_2x+b_2y+c_2z &= d_2 && \ldots (2) \qquad (\mathbf{d} \neq \mathbf{0}) \\ a_3x+b_3y+c_3z &= d_3 && \ldots (3) \end{aligned}$$

We multiply equation 1 by A_1, equation 2 by A_2, and equation 3 by A_3, and add.

Then,
$$\begin{aligned}&(a_1A_1+a_2A_2+a_3A_3)x\\ &\quad+(b_1A_1+b_2A_2+b_3A_3)y\\ &\qquad+(c_1A_1+c_2A_2+c_3A_3)z\\ &\qquad\quad= d_1A_1+d_2A_2+d_3A_3.\end{aligned}$$

The coefficient of x is just Δ (by column expansion) while those of y and z are zero (alien cofactors), so then

$$\Delta x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix} = D, \text{ say.}$$

Hence $\Delta x = D$, and similar expressions exist for y and z, enabling us to obtain (provided $\Delta \neq 0$) unique values for x, y and z.

This is not to be recommended as a method of solution in practical numerical problems, but it is important in the theory both of pure algebra and of numerical analysis as we shall show.

Cramer's method gives a very quick demonstration that when $\Delta = 0$ there is no unique solution in x, y, z; but, more than this, we see the two *types* of failure:

(*a*) If $D = 0$ as well as Δ, then *any* value of x will satisfy. (We then consider as a new problem the solution for y and z in terms of the coefficients and a chosen value of x.)

(*b*) If $\Delta = 0$ and $D \neq 0$, no solution is possible: the equations are inconsistent.

In numerical work we are also concerned with the situation called *ill-conditioning* of equations: if the value of Δ is small compared with the largest product of three of the terms a, b, c of which it is composed, then a slight error† in one of the elements may make a large percentage error in Δ and hence in any solution. (The geometrical picture is of three planes $\tilde{\pi}_1$, $\tilde{\pi}_2$ and $\tilde{\pi}_3$ which nearly meet in a line: if $\tilde{\pi}_1$ is rocked, then the point where it meets the line of intersection $\tilde{\pi}_2 \cap \tilde{\pi}_3$ may move a long way.

† The error could be in data, but it could also be in processing, e.g. in rounding off a coefficient to 5 significant figures.

16.4 *Adjoint matrix*

One of the properties of the cofactors of determinants turns out to be spectacularly useful. If we consider the matrix of cofactors of

$$\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix},$$

we have

$$\mathbf{N} = \begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{pmatrix}.$$

If we compute **MN** nothing very remarkable happens, but if we first reflect **N** about the main diagonal (called transposing, see section 14.4) and call the reflection **N**′, and then consider **MN**′, we get

$$\mathbf{MN}' = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{pmatrix}$$

$$= \begin{pmatrix} a_1A_1+b_1B_1+c_1C_1 & a_1A_2+b_1B_2+c_1C_2 & a_1A_3+b_1B_3+c_1C_3 \\ a_2A_1+b_2B_1+c_2C_1 & a_2A_2+b_2B_2+c_2C_2 & a_2A_3+b_2B_3+c_2C_3 \\ a_3A_1+b_3B_1+c_3C_1 & a_3A_2+b_3B_2+c_3C_2 & a_3A_3+b_3B_3+c_3C_3 \end{pmatrix}.$$

In the product matrix, all the terms on the main diagonal are det **M**, and all the terms of the main diagonal are zero (property of alien cofactors).

Hence
$$\mathbf{MN}' = \begin{pmatrix} \Delta & 0 & 0 \\ 0 & \Delta & 0 \\ 0 & 0 & \Delta \end{pmatrix}$$

$$= \Delta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \Delta\mathbf{I}.\dagger$$

This is a significant result, with **N**′ deserving a special name. We call it the 'adjoint matrix', written adj **M**, and the result is

$$\mathbf{M} \,.\, \text{adj}\, \mathbf{M} = \Delta\mathbf{I}. \qquad \dots (1)$$

This result has been proved only for 3×3 matrices, but it is perfectly general. We shall assume that it works for all square matrices without further proof.

† See appendix.

The importance of this result lies in the fact that it provides another method of finding the inverse matrix, provided that equation 1 may be divided by the number Δ, i.e. if $\Delta \neq 0$. Hence, if $\Delta \neq 0$, we find

$$\mathbf{M}\,.\left(\frac{1}{\Delta}\operatorname{adj}\mathbf{M}\right) = \mathbf{I}$$

$$\text{so that}\quad \frac{1}{\Delta}\operatorname{adj}\mathbf{M} = \mathbf{M}^{-1}.$$

Example

Find the inverse of $\mathbf{M} = \begin{pmatrix} 6 & 3 & 1 \\ 4 & 2 & -3 \\ 1 & 4 & 7 \end{pmatrix}$.

The cofactors corresponding to the various elements are:

1st row	26	−31	14
2nd row	−17	41	−21
3rd row	−11	22	0

Thus
$$\operatorname{adj}\mathbf{M} = \begin{pmatrix} 26 & -17 & -11 \\ -31 & 41 & 22 \\ 14 & -21 & 0 \end{pmatrix}.$$

Checking $\mathbf{M}\,.\operatorname{adj}\mathbf{M}$ to avoid mistakes,

$$\mathbf{M}\,.\operatorname{adj}\mathbf{M} = \begin{pmatrix} 6 & 3 & 1 \\ 4 & 2 & -3 \\ 1 & 4 & 7 \end{pmatrix}\begin{pmatrix} 26 & -17 & -11 \\ -31 & 41 & 22 \\ 14 & -21 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 156-93+14 & -102+123-21 & -66+66 \\ 104-62-42 & -68+82+63 & -44+44 \\ 26-124+98 & -17+164-147 & -11+88 \end{pmatrix}$$

$$= \begin{pmatrix} 77 & 0 & 0 \\ 0 & 77 & 0 \\ 0 & 0 & 77 \end{pmatrix}.$$

Hence
$$\mathbf{M}^{-1} = \tfrac{1}{77}\begin{pmatrix} 26 & -17 & -11 \\ -31 & 41 & 22 \\ 14 & -21 & 0 \end{pmatrix}.$$

The reader is warned against becoming too enthusiastic about the adjoint method of calculating the inverse matrix. It is good when calculating inverses of 2×2 or 3×3 matrices with simple coefficients, but

higher than that it is better to use the method of row operations. To illustrate this a table is given below of the number of arithmetic operations each method involves. The expressions for the last two cases give the order of magnitude only.

Size of matrix	*Adjoint method*	*Row method*
2×2	7	24
3×3	23	84
4×4	247	200
5×5	1609	390
100×100	10^{159}	3×10^{6}
$n\times n$	$en(n!)$	$3n^{3}$

It is seen that the adjoint method loses efficiency very rapidly indeed.

EXERCISE 16b

(1) Evaluate the inverse of the following matrix using first the adjoint method and then the row method:

$$\begin{pmatrix} 3 & -3 & -2 \\ 1 & 5 & 3 \\ 1 & 2 & 1 \end{pmatrix}$$

(2) Invert the following matrices, if possible, using either the adjoint method or the row method, whichever seems easier.

(*a*) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 4 & 7 & 2 \\ -3 & 1 & -7 \\ 2 & 4 & 1 \end{pmatrix}$ (*c*) $\begin{pmatrix} 2 & 3 & -1 \\ 4 & 2 & 0 \\ 0 & 4 & -2 \end{pmatrix}$

(*d*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 3 & 1 \end{pmatrix}$ (*e*) $\begin{pmatrix} 1 & 2 & 3 \\ -1 & 7 & -4 \\ 0 & 9 & -1 \end{pmatrix}$ (*f*) $\begin{pmatrix} 2 & 3 & -7 \\ 4 & 6 & 1 \\ -2 & 3 & 9 \end{pmatrix}$

(3) Show that the conic $x^2-2xy+2y^2-1 = 0$ may be written

$$(x\ y\ 1)\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (0)$$

or
$$\mathbf{x}'\,\mathbf{A}\,\mathbf{x} = 0.$$

Find the matrix $\operatorname{adj}\mathbf{A}$ and expand the equation

$$(l\ m\ 1)\operatorname{adj}\mathbf{A}\begin{pmatrix} l \\ m \\ 1 \end{pmatrix} = (0).$$

Find any solution l, m of this equation, and verify that the line $lx+my+1 = 0$ is a tangent to the original conic.

(4) The relationship between the stress vector **P** and the strain vector **e** when forces are applied normal to the faces of a rectangular specimen of material is

$$E\,\mathbf{e} = \begin{pmatrix} 1 & -\sigma & -\sigma \\ -\sigma & 1 & -\sigma \\ -\sigma & -\sigma & 1 \end{pmatrix} \mathbf{P}$$

where E and σ are constants known as 'Young's modulus' and 'Poisson's ratio' respectively. Find the inverse of the matrix and re-write the relation giving **P** in terms of **e**.

(5) Invert the matrix

$$\begin{pmatrix} \cos\theta\cos\phi & \sin\theta\cos\phi & -\sin\phi \\ \cos\theta\sin\phi & \sin\theta\sin\phi & \cos\phi \\ \sin\theta & -\cos\theta & 0 \end{pmatrix}.$$

Do you notice anything about your result?

EXERCISE 16c

The following questions are intended for further practice in inverting matrices.

(1) (*a*) $\begin{pmatrix} 1 & 11 & 3 \\ -4 & 5 & 1 \\ 2 & 3 & 1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 3 \\ -3 & 7 & -10 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$

(2) (*a*) $\begin{pmatrix} 2 & 4 & -1 \\ -4 & -7 & 5 \\ -4 & -5 & 12 \end{pmatrix}$ (*b*) $\begin{pmatrix} 2 & 3 & 1 \\ -6 & -8 & 2 \\ 14 & 19 & -1 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 4 & -3 \\ -6 & 5 & 2 \\ -3 & 17 & -7 \end{pmatrix}$

(3) (*a*) $\begin{pmatrix} 1 & -2 & -1 \\ -1 & 5 & 3 \\ 3 & -12 & -6 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & -2 & -3 \\ 2 & -3 & -8 \\ 3 & -4 & -12 \end{pmatrix}$

chapter 17

HOMOGENEOUS EQUATIONS NON-REGULAR EQUATIONS

Up to this point, in all our discussions of the triad of equations $a_1x+b_1y+c_1z = d_1$ etc.,
i.e.

$$\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{d},$$

we have stipulated (*i*) that the vector **d** is non-zero and (*ii*) that Δ ($=\det \mathbf{M}$) is non-zero. We shall now repair these omissions.

17.1 *Homogeneous equations*

We shall now consider the case $\mathbf{d} = \mathbf{0}$: the right-hand side coefficients of all our equations are zero. Such equations are described as *homogeneous* linear equations, since all terms are of the same degree (viz. degree *one* here) in the unknown quantities $x, y, z, \ldots$.

It is clear that $x = y = z = 0$ is a solution of such a triad of equations. The only question at issue is whether it is unique. The reader will not be surprised to discover that this is again decided by the value of Δ: the vanishing of Δ introduces as in previous cases one or more degrees of freedom† into the solutions. With the techniques which we now have available, this can be proved very neatly by either a matrix-wise or a vector-wise method:

Theorem: If $\det \mathbf{M} \neq 0$, then $\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0} \Leftrightarrow x = y = z = 0$.

The implication from right to left is clear by substitution, and we shall only be concerned with left to right.

† This means that x, y, z are expressible in terms of one parameter which can be given any value. Two degrees of freedom implies two parameters at choice e.g. $x = t+u$, $y = 2t-3u$, $z = u$.

Method (*a*): Since det $\mathbf{M} \neq 0$, $\mathbf{M}^{-1}$ exists.†

Then
$$\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0} \Rightarrow \mathbf{M}^{-1}\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{M}^{-1}\mathbf{0} = \mathbf{0}$$
$$\Rightarrow \mathbf{I}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$$
$$\Rightarrow x = y = z = 0.$$

Method (*b*): Since $\Delta \neq 0$, the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are linearly independent; i.e. no values x, y, z exist (not all zero) such that $x\mathbf{a}+y\mathbf{b}+z\mathbf{c} = \mathbf{0}$. Thus the zero values, which evidently satisfy, form the unique solution.

EXERCISE **17a**

(1) Find the value of λ such that the following triad of equations should have an infinity of solutions:

$$\begin{aligned} x+2y-\ z &= 0 \\ 5x+3y+2z &= 0 \\ 4x-\ y+\lambda z &= 0. \end{aligned}$$

Show that for this value of λ the ratio $x:y:z$ is the same for all non-zero solutions, and find it. Interpret the result geometrically.

(2) Find the values of λ such that the equations $3x+\lambda y+z = 0$, $x+y+z = 0$, $\lambda x+9y+4z = 0$ have an infinity of solutions, and find the ratio $x:y:z$ in each case.

(3) Give a general solution for the set of equations

$$\begin{aligned} x+\ 2y-\ z &= 0 \\ 3x-\ 4y+2z &= 0 \\ \lambda x-10y+5z &= 0, \end{aligned}$$

paying careful attention to any critical value(s) of λ.

(4) Three planes passing through the origin

$$\begin{aligned} 3x-2y+\ z &= 0 \\ x+\ y-2z &= 0 \\ \lambda x+7y-8z &= 0 \end{aligned}$$

intersect in a line. Determine the value of λ, and find the direction numbers of the line.

† $\mathbf{M}^{-1}$ is also unique, but this proof does not require it.

(5) The matrix $\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$ operating upon a certain vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is known to enlarge it, but not to change its direction, i.e.

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where λ is the enlargement factor. Rewrite this relationship as a set of homogeneous equations in x, y, and z and find the possible values of λ. For each value of λ, find the ratio $x:y:z$.

(6) Repeat question 5 using the matrix $\begin{pmatrix} 3 & 1 & -2 \\ 2 & 4 & -4 \\ 2 & 5 & -5 \end{pmatrix}$.

17.2 *Two homogeneous equations in three unknowns*

This can be regarded as a special case of the situation in which $\mathbf{d}$ and Δ are both zero. Given the equations

$$\begin{aligned} a_1x+b_1y+c_1z &= 0 \\ a_2x+b_2y+c_2z &= 0, \end{aligned}$$

we shall add the equation $0{\cdot}x+0{\cdot}y+0{\cdot}z = 0$ in order to make the number of equations up to three. Clearly the determinant of coefficients is zero.

An elementary solution could be carried out for x, y in terms of z by treating c_1z, c_2z (transferred to the right-hand side if preferred) as if they were numerical coefficients. We should obtain x/z and y/z uniquely, provided the determinant $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ were not zero. There is, however, a much neater way of proceeding, which has the advantage of generalising into higher orders of equations, as follows.

We have just seen that the problem is to find *ratios* of $x:y:z$. Thus if we can show a set of values, not all zero, to satisfy the equations, we have our solution. Now if we use the properties of alien cofactors of any determinant (whether Δ itself if zero or not), we have:

$$\begin{aligned} a_1 A_3+b_1B_3+c_1 C_3 &= 0 \\ a_2 A_3+b_2B_3+c_2 C_3 &= 0. \end{aligned}$$

Thus our required solution is $x:y:z = A_3:B_3:C_3$.

Notice that these are cofactors of the (zero) coefficients of our extra equation, but they do not themselves involve the zeros.

Our solution can be written:

$$x:y:z = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} : -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} : \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

This solution still holds good if one of these cofactors is zero; but not if two are zero, as the reader may verify.

The student should also verify that this rule holds good for the ratios in exercise 17a, nos. (1), (2) and (3). Notice that in nos. (1) and (3) the ratios are obtainable disregarding the λ-equation, provided we know that it is redundant (i.e. it gives no further information).

Redundancy of an equation can be considered here in another way, viz. as linear dependence, homogeneous equations being treated as *row*-vectors: then redundancy

$$\Leftrightarrow p\times(\text{equation I}) + q\times(\text{equation II}) + r\times(\text{equation III}) \equiv 0$$

for values p, q, r not all zero. Values p, q, r can be obtained by taking cofactors of any column, e.g. of the third column in exercise 17a, question (1), viz. $-17, 9, -7$.

The idea of linear dependence of the equations themselves is also useful when we have too many equations rather than too few: we shall not pursue it further here (but see exercise 17b (3)).

EXERCISE 17b

(1) (*a*) A plane passes through the origin and the points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$. Taking the equation of the plane as $lx+my+nz = 0$, solve for the ratio $l:m:n$.

(*b*) Write the condition that the tetrahedron OPP_1P_2 should have no volume, where P is (x, y, z), and compare your result with (*a*).

(2) (*a*) Given the vectors $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$, find the direction ratios l, m, n of a vector orthogonal to both of them.

(*b*) Comparing the result of (*a*) with that of the previous question, deduce a general result about the normal to a plane.

(3) Given that the following four planes meet in a point:

$$\begin{aligned} 2x+y-z &= 0 \\ 4x+3y+z &= 0 \\ x-y+2z &= 14 \\ x+y+z &= p \end{aligned}$$

find the value of p. (*Hint*: When the ratio $x:y:z$ is known, the third equation will determine them all.)

(4) The equations $ax^2+bx+c = 0$ and $bx^2+cx+a = 0$ have a common root α. Prove by eliminating α between the equations

$$\begin{aligned} a\alpha^2+b\alpha+c &= 0 \\ b\alpha^2+c\alpha+a &= 0 \end{aligned}$$

that the condition for this is

$$(ab-c^2)(ac-b^2) = (bc-a^2)^2.$$

(*Hint*: These equations could be solved in the form $\alpha^2:\alpha:1 = A:B:C$ i.e. like homogeneous equations.)

(5) By eliminating θ between the equations

$$\begin{aligned} a\sin\theta + \cos\theta + b &= 0 \\ \sin\theta + b\cos\theta + a &= 0, \end{aligned}$$

show that

$$(a-b^2)^2+(b-a^2)^2 = (ab-1)^2.$$

(*Hint*: As for question 4.)

17.3 We now consider what happens if $\Delta = 0$, $\mathbf{d} \neq 0$, starting with the case of two equations in two unknowns:

$$\begin{cases} px+qy = a \\ rx+sy = b \end{cases} \quad\text{or}\quad \begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The reader will recall that the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is an operator mapping the vector $\begin{pmatrix} x \\ y \end{pmatrix}$, but we remember that we can deduce what happens to $\begin{pmatrix} x \\ y \end{pmatrix}$ by seeing the effect on a basis.

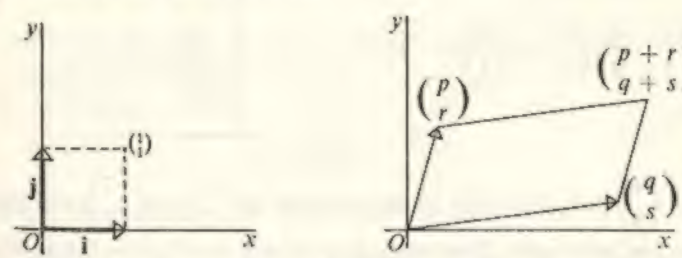
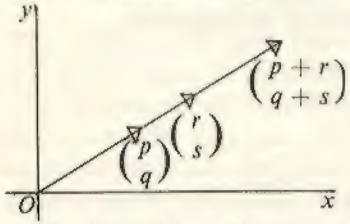


FIG. 63

The basis **i**, **j** is mapped into $\begin{pmatrix} p \\ r \end{pmatrix}$ and $\begin{pmatrix} q \\ s \end{pmatrix}$ such that the area of the parallelogram is $\Delta = ps - qr$.



But $\Delta = 0$, so the right-hand diagram collapses to look like Fig. 64.

FIG. 64

It is clear that the images of **i** and **j** no longer form a basis for the plane—in fact, $\begin{pmatrix} p \\ r \end{pmatrix}$ and $\begin{pmatrix} q \\ s \end{pmatrix}$ are linearly dependent and only span a space consisting of the line joining them to the origin. If the vector $\begin{pmatrix} a \\ b \end{pmatrix}$ lies on this line we shall be lucky and have solutions; if $\begin{pmatrix} a \\ b \end{pmatrix}$ does not lie on the line no solutions exist. We shall consider two simple examples, both with the same left-hand side:

(*a*) $\left.\begin{aligned} x+2y &= 5 \\ 2x+4y &= 2 \end{aligned}\right\}$

or $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

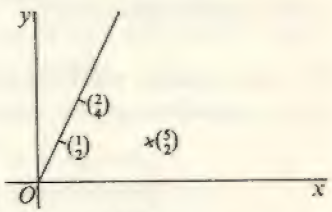


FIG. 65

Since $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ does *not* lie on the line joining $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ to the origin, i.e. it is not in the space spanned by $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$, there are *no* solutions.

(*b*) $\left.\begin{aligned} x+2y &= 4 \\ 2x+4y &= 8 \end{aligned}\right\}$

or $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$.

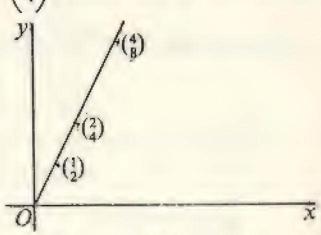


FIG. 66

Now $\begin{pmatrix} 4 \\ 8 \end{pmatrix}$ does lie in the space spanned by $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ and clearly we can find solutions. Two obvious ones are $x = 4, y = 0$ and $x = 0, y = 2$, but it is evident that we really have only one equation, $x+2y=4$,

and any values of x and y satisfying this will be a solution. We thus give x a particular value, t, say, and find y. Then,

$$x = t,\ y = \frac{4-t}{2}$$

is a solution for all values of t.

17.4 Leaving geometrical considerations aside, the routine solution of the equations of the previous section would be as follows:

(*a*) $\left.\begin{aligned} x+2y &= 5 \\ 2x+4y &= 2 \end{aligned}\right\}$

$\Leftrightarrow \left.\begin{aligned} x+2y &= 5 \\ 0 &= -8 \end{aligned}\right\} \quad r_2' = r_2 - 2r_1$

The two original equations are equivalent to a set of equations of which one is $0 = -8$, a statement which *no value of x and y* can make true. There is no solution to these equations and they are said to be INCONSISTENT.

(*b*) $\left.\begin{aligned} x+2y &= 4 \\ 2x+4y &= 8 \end{aligned}\right\}$

$\Leftrightarrow \left.\begin{aligned} x+2y &= 4 \\ 0 &= 0 \end{aligned}\right\} \quad r_2' = r_2 - 2r_1$

Now, the second equation gives no further information. It is true *whatever the values of x and y*. This (or the original) set of equations involved REDUNDANCY.† We now give x a particular value, t, and

$$x = t,\ y = \frac{4-t}{2}.$$

(*Note*: There is no reason why the particular value should be given to x. We might give it to y instead, so that $y = t$, $x = 4-2t$ is an equally acceptable solution. We also accept $x = 2-2t$, $y = 1+t$, and there are obviously countless alternatives.)

17.5 To show that the logic of a two-equation system can be quite exacting we consider the following example.

Given that

$$\begin{aligned} x-y &= -1 \\ mx-y &= a, \end{aligned}$$

to discuss the solutions for all possible values of m and a.

† The rows of coefficients, including the right-hand side, are like a set of linearly dependent row-vectors: one is a linear combination of the others.

The equations may be written

$$\begin{pmatrix} 1 & -1 \\ m & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ a \end{pmatrix}$$

and it follows immediately from the determinant that the equations are regular if $m \neq +1$.

Solving in the usual way gives

$$x = \frac{1+a}{m-1}, \quad y = \frac{m+a}{m-1} \quad (m \neq 1).$$

If $m = 1$, the equations are

$$\left.\begin{aligned} x-y &= -1 \\ x-y &= a \end{aligned}\right\}$$

and these are equivalent to

$$\left.\begin{aligned} x-y &= -1 \\ 0 &= a+1 \end{aligned}\right\} \quad r_2' = r_2 - r_1.$$

If $a \neq -1$, the equations are inconsistent.

If $a = -1$, the equations are consistent, the general solution being $x = t$, $y = t+1$.

The final solution is then:

(*a*) If $m \neq 1$, $c = \dfrac{1+a}{m-1}$, $y = \dfrac{m+a}{m-1}$.

(*b*) If $m = 1$ and $a = -1$, $x = t$, $y = t+1$.

(*c*) If $m = 1$, $a \neq -1$; no solution.

EXERCISE 17c

(1) Discuss the equations

$$\begin{aligned} x-ay &= 1 \\ x+2y &= b \end{aligned}$$

for all values of a and b.

(2) For which value of m will the equations

$$\begin{aligned} y-3x &= a \\ y-mx &= 2 \end{aligned}$$

be non-regular? Solve the equations in the non-regular case.

(3) Solve the equations

$$\begin{aligned} ax+3y &= 4 \\ bx-ay &= b \end{aligned}$$

considering all values of a and b.

(4) Discuss the solution of the equations

$$\begin{aligned} \lambda x+ y &= \mu \\ \mu x+\lambda y &= \mu^2 \end{aligned}$$

when $\lambda^2 = \mu$.

(5) Find the condition for the equations

$$\begin{aligned} ax+y &= b \\ x+by &= a \end{aligned}$$

to be non-regular, and solve them in this instance.

(6) Discuss the solution of the equations

$$\begin{aligned} y-mx &= 2, \\ y-x &= b \end{aligned}$$

for all possible values of m and b, giving interpretations and illustrations from co-ordinate geometry.

17.6 When we consider the 3×3 matrix of the set of equations

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

the basis vectors of the original space (viz. **i**, **j** and **k**) are transformed into

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \text{ and } \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

which we shall call **a**, **b** and **c** respectively. The reader will recall the remark in chapter 14, that the volume of the parallelepiped of which **a**, **b** and **c** are adjacent sides is Δ; and $\Delta = 0$ implies that this parallelepiped collapses into a plane (or possibly even a line). Alternatively, $\Delta = 0$ means that **a**, **b** and **c** are linearly dependent, which again means that the vectors **a**, **b** and **c** lie in a plane (or line) when attached to the origin. The vectors **a**, **b** and **c** no longer form a basis for the

vector 3-space, and now only span a lesser space; so, once again, as in the 2×2 equations, it is a matter of luck whether the right-hand side vector lies in that space or not.

We consider some examples with $\Delta = 0$.

(*a*) To discuss the equations:

$$\left.\begin{aligned} x+3y+4z &= 1\\ 3x-y+2z &= 1\\ 3x-11y-8z &= 1 \end{aligned}\right\}.$$

Now,

$$\left.\begin{aligned} x+3y+4z &= 1\\ 3x-y+2z &= 1\\ 3x-11y-8z &= 1 \end{aligned}\right\}$$

$$\Leftrightarrow \left.\begin{aligned} x+3y+4z &= 1\\ -10y-10z &= -2\\ -20y-20z &= -2 \end{aligned}\right\} \begin{aligned} &\\ r_2' &= r_2-3r_1\\ r_3' &= r_3-3r_1 \end{aligned}$$

$$\Leftrightarrow \left.\begin{aligned} x+3y+4z &= 1\\ -10y-10z &= -2\\ 0 &= 2 \end{aligned}\right\} \begin{aligned} &\\ &\\ r_3' &= r_3-2r_2 \end{aligned}$$

These equations are inconsistent. Clearly $\begin{pmatrix}1\\1\\1\end{pmatrix}$ does not lie in the space spanned by $\begin{pmatrix}1\\3\\3\end{pmatrix}$, $\begin{pmatrix}3\\-1\\-11\end{pmatrix}$ and $\begin{pmatrix}4\\2\\-8\end{pmatrix}$.

(*b*) To solve the equations:

$$\left.\begin{aligned} x-y+2z &= 2\\ 3x+2y-z &= 4\\ 5x+10y-11z &= 4 \end{aligned}\right\}.$$

Now,

$$\left.\begin{aligned} x-y+2z &= 2\\ 3x+2y-z &= 4\\ 5x+10y-11z &= 4 \end{aligned}\right\}$$

$$\Leftrightarrow \left.\begin{aligned} x-y+2z &= 2\\ 5y-7z &= -2\\ 15y-21z &= -6 \end{aligned}\right\} \begin{aligned} &\\ r_2' &= r_2-3r_1\\ r_3' &= r_3-5r_1 \end{aligned}$$

$$\left.\begin{aligned} x-y+2z &= 2\\ 5y-7z &= -2\\ 0 &= 0 \end{aligned}\right\} \begin{aligned} &\\ &\\ r_3' &= r_3-3r_2. \end{aligned}$$

These equations are consistent, with the third equation giving no extra information. Thus, putting $x = t$, we get

$$\left.\begin{aligned} -y+2z &= 2-t\\ 5y-7z &= -2 \end{aligned}\right\}$$

$$\Leftrightarrow \left.\begin{aligned} -y+2z &= 2-t\\ 3z &= 8-5t \end{aligned}\right\} \begin{aligned} &\\ r_2' &= r_2+5r_1 \end{aligned}$$

$$\Leftrightarrow x = t,\ z = \frac{8-5t}{3},\ y = \frac{10-7t}{3}$$

(*c*) To find the value of λ which makes the set of equations

$$\begin{aligned} x+2y+3z &= 10\\ 4x-y+\lambda z &= a\\ 2x-5y-7z &= -11 \end{aligned}$$

non-regular, and discuss the solution for that value.

The equations are non-regular when $\Delta = 0$ and

$$\begin{aligned} \Delta = \begin{vmatrix} 1 & 2 & 3\\ 4 & -1 & \lambda\\ 2 & -5 & -7 \end{vmatrix} &= (7+5\lambda)-2(-28-2\lambda)+3(-20+2)\\ &= 7+56-54+9\lambda\\ &= 9+9\lambda. \end{aligned}$$

Then $\Delta = 0 \Leftrightarrow \lambda = -1$.

When $\lambda = -1$, the equations become:

$$\left.\begin{aligned} x+2y+3z &= 10\\ 4x-y-z &= a\\ 2x-5y-7z &= -11 \end{aligned}\right\}$$

$$\Leftrightarrow \left.\begin{aligned} x+2y+3z &= 10\\ -9y-13z &= a-40\\ -9y-13z &= -31 \end{aligned}\right\} \begin{aligned} &\\ r_2' &= r_2-4r_1\\ r_3' &= r_3-2r_1. \end{aligned}$$

The equations are consistent if $a-40 = -31 \Leftrightarrow a = 9$. The equations become:

$$\begin{aligned} 2y+3z &= 10\\ 9y+13z &= 31. \end{aligned}$$

Putting $x = t$, we have

$$\left.\begin{aligned} x+2y+3z &= 10-t\\ 9y+13z &= 31 \end{aligned}\right\}$$

$$\Leftrightarrow \begin{aligned} \left.\begin{aligned} 2y+3z &= 10-t\\ -z &= -28+9t \end{aligned}\right\} &\quad r_2' = 2r_2-9r_1\\ 2y = 10-t-84+27t. & \end{aligned}$$

The solution is then $x = t$, $y = 13t-37$, $z = 28-9t$, if $a = 9$.

EXERCISE **17d**

(1) Discuss the equations

$$\begin{cases} 2x+\ y-3z = 2 \\ 3x-2y+\ z = 8 \\ \ x+4y-7z = k \end{cases}$$

for the cases $k = 0$, $k = -4$. (In the case for which there is a degree of freedom put $z = 7t$.)

(2) By treating a linear equation of form $ax+by = c$ as a row-vector $(a, b, -c)$, or otherwise, find the value or values of λ for which the following lines are concurrent:

$$\lambda x+2y = 12,\ x-\lambda y = 9,\ 2x+3y = 1.$$

Find the point of concurrency in each case.

(3) Investigate the solutions of the equations

$$\begin{aligned} 3ax-6y+18z &= b \\ 4x+5y-\ z &= -2 \\ x-7y-\ z &= 3 \end{aligned}$$

in the cases (*i*) $a = 3$, $b = 21$; (*ii*) $a = -17$, $b = 1$.

(*Maths Tripos Pt.* 1, 1949)

(4) For what value of λ will the following set of equations have more than one solution?

$$\begin{aligned} \lambda x+\ y-\ z &= 0 \\ 3x+4y-3z &= 0 \\ 5x-3y+2z &= 0 \end{aligned}$$

(5) Examine the solutions, in the non-regular case only, of the equations

$$\begin{aligned} 2x+\lambda y-6z &= 2, \\ \lambda x-3y+2z &= -1, \\ 3x-2y-4z &= \lambda. \end{aligned}$$

(6) Find the general solution of the system of equations

$$\begin{aligned} x+\ y-az &= b, \\ 3x-2y-\ z &= 1, \\ 4x-3y-\ z &= 2, \end{aligned}$$

in each of the three special cases:
(*i*) $a = 1, b = 9$; (*ii*) $a = 2, b = -3$; (*iii*) $a = 2, b = 0$.

(*Cambridge Scholarship*)

(7) Prove that, if $a \neq 5$, the equations

$$\begin{aligned} x+ay-z &= 1, \\ 2x+\ y+z &= b, \\ x-\ y+z &= 2 \end{aligned}$$

have a unique solution.

If $a = 5$, classify the values of b such that (*i*) the equations have no solutions, (*ii*) the equations have an infinite number of solutions.

(*Oxford Scholarship*)

(8) Solve for x, y, z, the three simultaneous equations

$$\begin{aligned} ax+3y+2z &= b, \\ 5x+4y+3z &= 1, \\ x+2y+\ z &= b^2 \end{aligned}$$

for the particular cases obtained when $a = 3$ for varying values of b.

(*Cambridge Scholarship*)

(9) Factorise the determinant

$$\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}.$$

Given that a, b, and c are real and not all equal and that $a+b+c \neq 0$, solve

$$\begin{aligned} ax+by+cz &= 1 \\ cx+ay+bz &= 0 \\ bx+cy+az &= 0. \end{aligned}$$

What happens when $a+b+c = 0$?

(*Cambridge Scholarship*)

(10) Prove that, if the simultaneous equations

$$\begin{aligned} 3x+ky+2z &= \lambda x, \\ kx+3y+2z &= \lambda y, \\ 2x+2y+\ z &= \lambda z \end{aligned}$$

have a solution in which x, y, z are not all zero, then

$$(1-\lambda)k^2-8k+(\lambda+1)(\lambda-3)(\lambda-5) = 0.$$

When this condition is satisfied, find formulae for the most general solutions in the two cases (*i*) $\lambda = 1$, (*ii*) $\lambda = 3$.

(*Cambridge Scholarship*)

chapter 18

SOME FURTHER PROPERTIES OF MATRICES

18.1 *Summary of properties developed*

Before going further we take the opportunity to restate those properties of matrices so far developed.

Matrices have been defined, together with a law of combination, viz. multiplication (provided they are conformable for it) and under this operation they are associative.

To each square matrix there corresponds a determinant; this may be considered as a mapping from the set of matrices on to the set of real numbers. If the determinant is zero, the matrix is said to be singular and has no inverse; if the determinant is non-zero, the matrix is non-singular and has a unique inverse.

We now consider two other algebraic properties, which will be required in subsequent work.

18.2 *Inverse of a product*

If we have two square non-singular matrices $\mathbf{A}$ and $\mathbf{B}$, we require to find a matrix $\mathbf{C}$ such that $(\mathbf{AB})\mathbf{C} = \mathbf{I}$ (i.e. an inverse for $\mathbf{AB}$).

Remembering that under multiplication matrices are associative we may omit the brackets and write $\mathbf{ABC} = \mathbf{I}$. Then multiplying by $\mathbf{A}^{-1}$ which exists (uniquely) since $\mathbf{A}$ is non-singular:

$$
\begin{aligned}
& \mathbf{A}^{-1}\mathbf{ABC} = \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1} \\
\Leftrightarrow \quad & \mathbf{IBC} = \mathbf{A}^{-1} \\
\Leftrightarrow \quad & \mathbf{BC} = \mathbf{A}^{-1} \\
\Leftrightarrow \quad & \mathbf{B}^{-1}\mathbf{BC} = \mathbf{B}^{-1}\mathbf{A}^{-1} \text{ (since } \mathbf{B} \text{ is non-singular)} \\
\Leftrightarrow \quad & \mathbf{IC} = \mathbf{B}^{-1}\mathbf{A}^{-1} \\
\Leftrightarrow \quad & \mathbf{C} = \mathbf{B}^{-1}\mathbf{A}^{-1}.
\end{aligned}
$$

We also see that $\mathbf{C}$ has the property that

$$\mathbf{CAB} = \mathbf{I}$$

for
$$
\begin{aligned}
\mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{AB} &= \mathbf{B}^{-1}\mathbf{IB} \\
&= \mathbf{B}^{-1}\mathbf{B} \\
&= \mathbf{I}.
\end{aligned}
$$

Thus, since $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = (\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{I}$, so that $\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the inverse of $\mathbf{AB}$.

Geometrically, the following mapping diagram is self-explanatory and is left without comment:

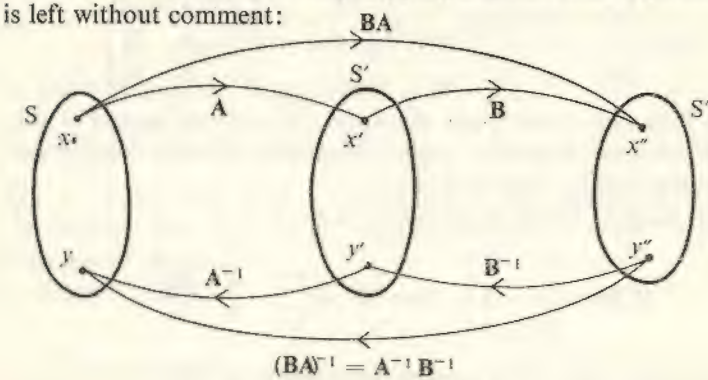


FIG. 67

EXERCISE **18a**

Reference to e.g. $\mathbf{X}^{-1}$ in this exercise may be taken to imply $\mathbf{X}$ to be non-singular.

(1) For the matrices $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$ write down $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$. Evaluate $\mathbf{AB}$ and verify that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

(2) Prove that $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

(3) The four matrices $\mathbf{I}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{AB}(=\mathbf{BA})$ form a group under multiplication in which $\mathbf{I}$ is the identity and every element is its own inverse. Show that if $\mathbf{T}$ is *any* chosen member of this group, then the transformation $\mathbf{X} \to \mathbf{TXT}^{-1}$ carried out over the whole group (i.e. by letting $\mathbf{X}$ be every member in turn) gives the same group again.

If the reader has difficulty in doing this in the abstract, it is possible to consider the case in which $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(4) Show that any 3×3 non-singular matrix $\mathbf{M}$ can be expressed as the product $\mathbf{LU}$ of a lower and an upper triangular matrix, in which one of these has all its main diagonal elements unity. (See Choleski's method: chapter 13.) Show that this opens up a method of finding $\mathbf{M}^{-1}$. Comment on the form of $\mathbf{L}^{-1}$ and $\mathbf{U}^{-1}$.

(5) If $\mathrm{T}(x) = x'$ where $x' = \dfrac{ax+b}{cx+d}$ find the condition for the inverse transformation to exist. If $\mathrm{T}_1(x) = \dfrac{x+1}{x-1}$ and $\mathrm{T}_2(x) = \dfrac{x+p}{p-x}$, find the form of $\mathrm{T}_2\mathrm{T}_1$ and its inverse.

18.3 *Transposing matrices*

We have already remarked that the matrix

$$\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \text{ is not the same as } \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

although their determinants are the same. The second matrix is said to be the *transpose* of the first, and is denoted by a prime. The second matrix would thus be written $\mathbf{M}'$.

There is no need for $\mathbf{M}$ to be square; e.g.

$$\text{if } \mathbf{M} = \begin{pmatrix} 3 & 2 \\ 4 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{then } \mathbf{M}' = \begin{pmatrix} 3 & 4 & 0 \\ 2 & 1 & 3 \end{pmatrix}.$$

As before, the rows of $\mathbf{M}$ become the columns of $\mathbf{M}'$.

In the applications of matrices, we often need to transpose a product of matrices and find the result $(\mathbf{AB})'$—the reader should experiment with simple matrices, preferably non-square, to find the rule for himself.

We suppose that $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times p$, and think of $\mathbf{A}$ as being made up of m row-vectors, $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_m$, and $\mathbf{B}$ as being made up of p column-vectors, $\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_p$. The element in the ith row and jth column of the product $\mathbf{AB}$, is then the inner product of the vectors $\mathbf{r}_i$ and $\mathbf{c}_j$. Thus the element in the jth row and ith column of $(\mathbf{AB})'$ is also $\mathbf{r}_i \,.\, \mathbf{c}_j$ which equals $\mathbf{c}_j \,.\, \mathbf{r}_i$. But this comes from a matrix in which the **c**-vectors are rows, multiplied by a matrix in which the **r**-vectors are columns, i.e. $\mathbf{B}'\mathbf{A}'$.

Thus, $$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'.$$

EXERCISE 18b

(1) For the two matrices $\mathbf{A} = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -3 & -5 \\ 1 & 1 \end{pmatrix}$ evaluate $\mathbf{A}'\mathbf{B}'$, $\mathbf{B}'\mathbf{A}'$, $(\mathbf{AB})'$, and $(\mathbf{BA})'$. Verify that $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ and $(\mathbf{BA})' = \mathbf{A}'\mathbf{B}'$.

(2) Show that whatever the shape of $\mathbf{A}$, it is always possible to compute the products $\mathbf{AA}'$ and $\mathbf{A}'\mathbf{A}$. Evaluate these products when

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 & 1 \\ 2 & -1 & 3 & -1 \\ 1 & 4 & 1 & -3 \end{pmatrix}$$

and make an observation about the numbers in each of these matrices. Is this coincidence?

(3) If $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 1 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ verify that $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

(4) Show that for any matrix $(\mathbf{M}')' = \mathbf{M}$.

(5) Show that for the matrix $\mathbf{M} = \begin{pmatrix} 3 & 1 \\ 9 & 4 \end{pmatrix}$ it is true that $(\mathbf{M}')^{-1} = (\mathbf{M}^{-1})'$. Prove that $\mathbf{X}'\mathbf{M} = \mathbf{I} \Leftrightarrow \mathbf{M}'\mathbf{X} = \mathbf{I}$. By putting $\mathbf{X} = (\mathbf{M}^{-1})'$ and showing that $\mathbf{X} = (\mathbf{M}')^{-1}$ prove that $(\mathbf{M}')^{-1} = (\mathbf{M}^{-1})'$ for any non-singular matrix.

(6) Simplify $(\mathbf{AB}'\mathbf{C}')'$.

18.4 *Symmetric and skew-symmetric matrices*

It sometimes happens (see exercise 18b, no. 2) that a matrix $\mathbf{A}$ may be equal to its own transpose, i.e. $\mathbf{A} = \mathbf{A}'$. If this is so the matrix is said to be *symmetric*; if, however, $\mathbf{A} = -\mathbf{A}'$, $\mathbf{A}$ is said to be *skew-symmetric*.

Examples of symmetric and skew-symmetric matrices are:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 3 & 0 \\ 2 & 0 & 4 \end{pmatrix} \quad \textit{Symmetric}$$

and $$\mathbf{B} = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix} \quad \textit{Skew-symmetric.}$$

A numerical matrix indicates its nature by inspection; but if we wish to test the symmetry of any abstract matrix $\mathbf{M}$, we attempt formally to prove that $\mathbf{M} = \mathbf{M}'$. As a worked example we show that $\mathbf{AA}'$ is always symmetric (and so is $\mathbf{A}'\mathbf{A}$).

If we put $\mathbf{AA}' = \mathbf{M}$, then we shall show that $\mathbf{M} = \mathbf{M}'$. We have already seen that, if $\mathbf{M} = \mathbf{AB}$, then $\mathbf{M}' = \mathbf{B}'\mathbf{A}'$. Putting $\mathbf{B} = \mathbf{A}'$ we have:

$$\begin{aligned} \mathbf{M}' &= (\mathbf{AA}')' \\ &= (\mathbf{A}')'\mathbf{A}' \\ &= \mathbf{AA}' \\ &= \mathbf{M} \text{ as required.} \end{aligned}$$

EXERCISE 18c

(1) Show that if **A** and **B** are both symmetric, **AB** is not necessarily symmetric unless a certain algebraic condition is satisfied. What is this condition?

(2) Given (*i*) a *diagonal* matrix **D**, i.e. all the terms off the main diagonal are zero, (*ii*) a square matrix **P** conformable for multiplication with **D**, and (*iii*) $\mathbf{M} = \mathbf{P'DP}$, prove that **M** is symmetric.

(3) If **A** is a skew-symmetric matrix, prove that $\mathbf{A}^2$ is symmetric. What about $\mathbf{A}^m$?

(4) If $\mathbf{l} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, show that $\mathbf{l'x}$ is symmetric. Is $\mathbf{lx'}$?

18.5 *Area and volume properties of matrices*

We have seen that the determinant of a 2×2 matrix is the area of the parallelogram into which the unit square is mapped.

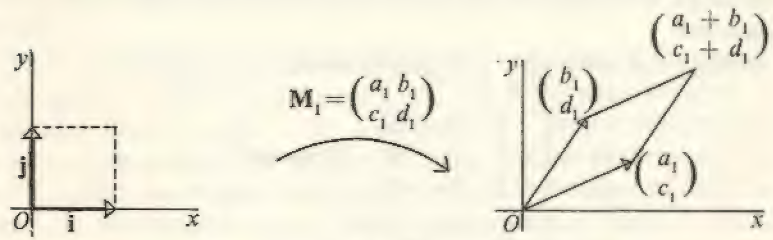


FIG. 68

The area of the parallelogram is $\det \mathbf{M}_1 = a_1d_1 - b_1c_1$, and this may be considered as an enlarging factor. Thought of in this way, it seems natural that when one matrix mapping $\mathbf{M}_1$, is followed by another, $\mathbf{M}_2$, then the enlargement factors $\det \mathbf{M}_1$ and $\det \mathbf{M}_2$ are multiplied. In fact, what we are saying is that, intuitively,

$$\det(\mathbf{M}_2\mathbf{M}_1) = \det \mathbf{M}_2 \det \mathbf{M}_1.$$

This is quite easy to verify when $\mathbf{M}_1$ and $\mathbf{M}_2$ are 2×2, but the verification when they are 3×3 makes it worth while looking for a different method of proof.

We go on to establish a proof that $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$, but before we start, the reader should work 18d to establish two necessary lemmas.

EXERCISE 18d

(Both these questions are lemmas for use in section 18.6.)

(1) By considering

$$\mathbf{E}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_2 = \begin{pmatrix} 1 & \lambda & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_3 = \begin{pmatrix} \mu & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

show that the inverse of an elementary matrix is itself elementary.

(2) For each of the elementary matrices **E** in question 1, consider the product **EM**, where

$$\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix}$$

and show that, in each case, $\det \mathbf{EM} = \det \mathbf{E} \det \mathbf{M}$.

18.6 *Determinant of a product*

We require to prove that $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$, and we first assume that **A** is non-singular. When we calculated the inverse of **A** using row-operations, we used a method which showed that $\mathbf{E}_k\mathbf{E}_{k-1} \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$, where the matrices **E** are all elementary. But the inverse of an elementary matrix is itself elementary, so that if we write the inverse of $\mathbf{E}_i$ as $\mathbf{F}_i$, we see that

$$\mathbf{A} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3 \ldots \mathbf{F}_{k-1}\mathbf{F}_k,$$

and

$$\mathbf{AB} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3 \ldots \mathbf{F}_{k-1}\mathbf{F}_k\mathbf{B}.$$

We now use the lemma, proved in exercise 18d, for

$$\begin{aligned} \det \mathbf{AB} &= \det(\mathbf{F}_1\mathbf{F}_2\mathbf{F}_3 \ldots \mathbf{F}_{k-1}\mathbf{F}_k\mathbf{B}) \\ &= \det \mathbf{F}_1 \det(\mathbf{F}_2\mathbf{F}_3 \ldots \mathbf{F}_{k-1}\mathbf{F}_k\mathbf{B}) \\ &= \det \mathbf{F}_1 \det \mathbf{F}_2 \det(\mathbf{F}_3 \ldots \mathbf{F}_{k-1}\mathbf{F}_k\mathbf{B}) \\ &= \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &= \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &= \det \mathbf{F}_1 \det \mathbf{F}_2 \ldots \det \mathbf{F}_k \det \mathbf{B}. \end{aligned}$$

But, $\det \mathbf{F}_1 \det \mathbf{F}_2 \ldots \det \mathbf{F}_k = \det \mathbf{A}$.

Hence, $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$ when **A** is non-singular.

When **A** is singular this method of proof breaks down, for **A** can no longer be reduced to the form $\mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$.

However, by using elementary matrices, we can reduce **A** to the matrix **R**, where **R** has the bottom row (at least) with only zeros.

Hence $\mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{R}$, and $\mathbf{A} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3\ldots\mathbf{F}_{k-1}\mathbf{F}_k\mathbf{R}$.

If we consider the product **RB**, we see by direct multiplication that, as the bottom row of **R** consists of zeros, the bottom row of **RB** will also consist of zeros, so that det **RB** = 0.

$$
\begin{aligned}
\text{Thus,}\quad \det \mathbf{AB} &= \det (\mathbf{F}_1\mathbf{F}_2\mathbf{F}_3\ldots\mathbf{F}_{k-1}\mathbf{F}_k\mathbf{RB})\\
&= \det \mathbf{F}_1 \det (\mathbf{F}_2\mathbf{F}_3\ldots\mathbf{F}_{k-1}\mathbf{F}_k\mathbf{RB})\\
&= \det \mathbf{F}_1 \det \mathbf{F}_2 \det (\mathbf{F}_3\ldots\mathbf{F}_{k-1}\mathbf{F}_k\mathbf{RB})\\
&= \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&= \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&= \det \mathbf{F}_1 \det \mathbf{F}_2 \ldots \det \mathbf{F}_k \det (\mathbf{RB})\\
&= 0.
\end{aligned}
$$

Now, since **A** is singular, det **A** = 0, thus:

$$\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}.$$

Thus det **AB** = det **A** det **B** in all cases.

EXERCISE 18e

(1) Verify that det **AB** = det **A** det **B** when

$$(i)\ \mathbf{A} = \begin{pmatrix} 2 & -1 \\ 6 & 3 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

$$(ii)\ \mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}.$$

(2) Prove that the product of two non-singular matrices is non-singular.

(3) Prove that det **AB** = det **BA** when **A** and **B** are square. Test the result when **A** is 2×3 and **B** is 3×2 by taking particular cases.

(4) Prove that $\det (\mathbf{A}^{-1}) = 1/\det \mathbf{A}$.

(5) (The reader should study again the meaning of adj **M**.) Show that, if **M** is a 3×3 matrix, $\det (\text{adj } \mathbf{M}) = m^2$, where $\det \mathbf{M} = m$. Prove also that $\text{adj } (\text{adj } \mathbf{M}) = m\mathbf{M}$.

(6) Prove that, if **A** and **B** are square and **AB** is singular, then either **A** or **B** is singular (or both).

(7) The general 2×2 matrix $\mathbf{M} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ can be made to satisfy a number of conditions, expressible in terms of its elements. State these conditions in the following cases:

(*a*) **Ai** = **i**. Express this also in geometrical terms.

(*b*) **B** transforms the unit square into a figure of unit area.

(*c*) **C** performs an overall linear magnification k while preserving all directions.

(*d*) **R** effects a pure rotation.

(*e*) **H** satisfies both conditions (*a*) and (*b*).

Of which type is $\mathbf{K} = \begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}$? What does the matrix $\mathbf{R}^{-1}\mathbf{KR}$ do to rectangle *OACB*, where $\mathbf{a} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$, $\mathbf{c} = \mathbf{a}+\mathbf{b}$ and $\mathbf{R} = \begin{pmatrix} 4/5 & 3/5 \\ -3/5 & 4/5 \end{pmatrix}$?

(8) Show that 2×2 matrices with unit determinant form a group G under multiplication. A subset of this group is the set $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$: show that this is also a sub-group, illustrating your answer geometrically. Name two different types of sub-group of G.

(9) Show that

$$\begin{vmatrix} b^2+c^2+1 & c^2+1 & b^2+1 & b+c \\ c^2+1 & c^2+a^2+1 & a^2+1 & c+a \\ b^2+1 & a^2+1 & a^2+b^2+1 & a+b \\ b+c & c+a & a+b & 3 \end{vmatrix}$$

is the square of a certain determinant, and hence obtain its value.

(*Cambridge Scholarship, adapted*)

(10) By writing the determinant

$$\begin{vmatrix} S_0 & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{vmatrix}$$

where $S_k = a^k+b^k+c^k$, as the product of two determinants, or otherwise, show that it is equal to $(b-c)^2(c-a)^2(a-b)^2$.

Find also the value of the determinant

$$\begin{vmatrix} S_0 & S_2 & S_4 \\ S_1 & S_3 & S_5 \\ S_2 & S_4 & S_6 \end{vmatrix}.$$

(*Math. Tripos*)

18.7 *Preservation of angle and of magnitude of vectors by a matrix mapping*

An important type of transformation is one for which the angle between any two vectors before transforming is the same as the angle after transforming. Since an enlargement has this effect trivially and we wish to look rather deeper, we will discuss only those matrices which preserve angle without enlargement.

For 2×2 matrices operating upon vectors in a plane, it is clear that a rotation is an example of this type of transformation, but no clear picture emerges in three dimensions.

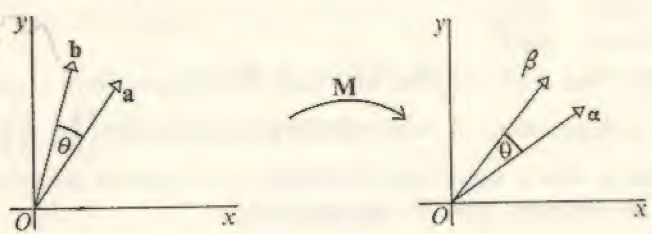


FIG. 69

We suppose that the transforming matrix is $\mathbf{M}$ and try to establish what underlying property $\mathbf{M}$ has in order to perform such a mapping.

Consider a 3×3 matrix $\mathbf{M}$ operating on two unit 3-vectors $\mathbf{a}$ and $\mathbf{b}$, or $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ respectively. Suppose that $\mathbf{Ma} = \boldsymbol{\alpha}$ and $\mathbf{Mb} = \boldsymbol{\beta}$.

The condition that no enlargement takes place, i.e. $\mathbf{a}$ and $\boldsymbol{\alpha}$ are both unit vectors, is expressed in vector language by taking the scalar product of $\mathbf{a}$ with itself, i.e. $\mathbf{a} \cdot \mathbf{a} = 1$. In matrix notation this is written $\mathbf{a}'\mathbf{a} = (a_1 a_2 a_3)\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (a_1^2 + a_2^2 + a_3^2) = (1)$.

The dash for the transpose might be said to take the place of the dot. Also we know that $\boldsymbol{\alpha}'\boldsymbol{\alpha} = (1)$, for $\boldsymbol{\alpha}$ is also a unit vector.

But since $\boldsymbol{\alpha} = \mathbf{Ma}$, $\boldsymbol{\alpha}' = \mathbf{a}'\mathbf{M}'$,

$$\boldsymbol{\alpha}'\boldsymbol{\alpha} = \mathbf{a}'\mathbf{M}'\mathbf{Ma}$$

and we now have the condition that

$$\mathbf{a}'\mathbf{a} = \mathbf{a}'\mathbf{M}'\mathbf{Ma}. \qquad \ldots \text{condition (1)}$$

The second condition concerns angles, and we find the angle between $\mathbf{a}$ and $\mathbf{b}$ also by scalar product methods. Here, remembering that $|\mathbf{a}|$ and $|\mathbf{b}|$ are both unity, the angle θ between $\mathbf{a}$ and $\mathbf{b}$ is given by $\mathbf{a}'\mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta = \cos\theta$ and we require that θ is also the angle between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Hence, $\boldsymbol{\alpha}'\boldsymbol{\beta} = \mathbf{a}'\mathbf{b}$,

and so $\mathbf{a}'\mathbf{M}'\mathbf{Mb} = \mathbf{a}'\mathbf{b} \qquad \ldots \text{condition (2)}$

Conditions (1) and (2) are clearly both satisfied if $\mathbf{M}'\mathbf{M} = \mathbf{I}$, so that a matrix $\mathbf{M}$, having this algebraic property, certainly does what we want geometrically.

The proof that $\mathbf{M}'\mathbf{M} = \mathbf{I}$ is *necessary* (so far we have only shown it is sufficient) is left to the reader as an exercise—exercise 18f, (4).

A matrix $\mathbf{M}$ with the property that $\mathbf{M}'\mathbf{M} = \mathbf{I}$ is said to be *orthogonal*. We see immediately that $\mathbf{M}' = \mathbf{M}^{-1}$, so that we have a very easy method of writing its inverse. Generally, finding an inverse can be laborious, so that to be able merely to transpose for the inverse makes it worth while algebraically to use orthogonal matrices whenever possible.

Examples of orthogonal matrices are:

(*a*)
$$\mathbf{A} = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}.$$

For this matrix
$$\mathbf{A}' = \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}$$

and
$$\mathbf{A}'\mathbf{A} = \begin{pmatrix} 9/25+16/25 & -12/25+12/25 \\ -12/25+12/25 & 16/25+9/25 \end{pmatrix} = \mathbf{I}.$$

That this matrix represents a rotation may easily be seen, for it is of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

which is itself orthogonal.

(*b*)
$$\mathbf{B} = \begin{pmatrix} 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix}.$$

The reader may easily verify that $\mathbf{B'B} = \mathbf{I}$, but it is no longer apparent what $\mathbf{B}$ represents.

EXERCISE 18f

(1) By finding the images of $\mathbf{i}$ and $\mathbf{j}$ write down the 2×2 matrix $\mathbf{M}$ which represents a reflection in the line $y = x \tan \alpha$. Test this matrix for orthogonality.

(2) Show that if the matrix $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is orthogonal, then the columns considered as vectors must be orthogonal. Does this also hold true for the matrix $\mathbf{N} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ when $\mathbf{N}$ is orthogonal?

(3) If $\mathbf{M}$ is the matrix $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ find the two values of λ for which the equation $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ has solutions in $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$. For each of these values of λ, find the corresponding direction given by the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ and find a unit vector in that direction. Make up a matrix with these vectors as columns, and show that it is orthogonal. If this matrix is $\mathbf{P} = \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix}$, show that $\mathbf{M} = \mathbf{P\,D\,P}^{-1}$ where $\mathbf{D}$ is the matrix $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

(4) To show that $\mathbf{M'M} = \mathbf{I}$ is necessary to preserve angles without enlargement, consider the images of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ and impose the condition that these images are mutually perpendicular. Show how this leads to $\mathbf{M'M} = \mathbf{I}$.

(5) If $\mathbf{A}^{-1}\mathbf{BA}$ is symmetric, and $\mathbf{A}$ is orthogonal, prove that $\mathbf{B}$ is symmetric: can you make any deduction about $\mathbf{B}$ if $\mathbf{A}^{-1}\mathbf{BA}$ is skew-symmetric?

(6) Prove that, if the matrix $\mathbf{M}$ is orthogonal, where
$$\mathbf{M} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

then the columns of $\mathbf{M}$ considered as vectors form an orthogonal basis; i.e. they form a basis in which every base-vector is a unit vector and orthogonal to every other base-vector.

18.8 *Addition of matrices*

Suppose we have two matrices, $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix}$, and we operate with each of these matrices on the vector $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$. We find that
$$\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x+y \\ 2x+4y \end{pmatrix} \text{ and } \mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x-3y \\ x+2y \end{pmatrix}.$$

If we now add these vectors, we see
$$\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix}+\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x+y \\ 2x+4y \end{pmatrix}+\begin{pmatrix} 2x-3y \\ x+2y \end{pmatrix} = \begin{pmatrix} 5x-2y \\ 3x+6y \end{pmatrix}.$$

We ask the question, is there a single mapping which gives this result? The answer clearly is yes, viz. $\begin{pmatrix} 5 & -2 \\ 3 & 6 \end{pmatrix}$.

In view of the construction of this matrix, we define it to be the matrix *sum* of $\mathbf{A}$ and $\mathbf{B}$.

Thus,
$$\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}+\begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & 6 \end{pmatrix}.$$

A very simple generalisation leads us to define addition of matrices (provided they are the same size) by adding corresponding terms, e.g.
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}+\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} a+p & b+q \\ c+r & d+s \end{pmatrix}.$$

Immediately the question arises, how does this operation fit into the pattern of multiplication that we have already? To show this, we prove the distributive rule, viz. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

To prove this, we suppose that the matrix $\mathbf{A}$ is made up of n row-vectors, $\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_n$, and that $\mathbf{B}$ and $\mathbf{C}$ are each made up of n column-vectors, $\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_n$ and $\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_n$. We consider the elements in the first row and first column of the matrices $\mathbf{A}(\mathbf{B}+\mathbf{C})$ and $\mathbf{AB}+\mathbf{AC}$.

The result for $\mathbf{A}(\mathbf{B}+\mathbf{C})$ will be $\mathbf{a}_1 . (\mathbf{b}_1+\mathbf{c}_1)$ and for $\mathbf{AB}+\mathbf{AC}$ will be $\mathbf{a}_1 . \mathbf{b}_1+\mathbf{a}_1 . \mathbf{c}_1$. But these two expressions are equal (see chapter 7), and so these elements are equal.

Similarly the elements in the ith row and jth column of $\mathbf{A}(\mathbf{B}+\mathbf{C})$ and $\mathbf{AB}+\mathbf{AC}$ are respectively:

$$\mathbf{a}_i . (\mathbf{b}_j+\mathbf{c}_j) \text{ and } \mathbf{a}_i . \mathbf{b}_j+\mathbf{a}_i . \mathbf{c}_j.$$

These are clearly equal, establishing the distributive rule

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}.\dagger$$

EXERCISE **18g**

(1) Show that vector addition is a special case of matrix addition.

(2) Verify the distributive law, $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$, when

$$\mathbf{A} = (2 \quad 1),\ \mathbf{B} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix},\text{ and } \mathbf{C} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}.$$

(3) Show that the matrix $\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$ satisfies the matrix equation $\mathbf{A}^3-6\mathbf{A}^2+11\mathbf{A}-6\mathbf{I} = \mathbf{Z}$, where $\mathbf{Z}$ is the 3×3 zero matrix. Is it permissible to factorise this expression in the usual way to find three matrices whose product is the zero matrix?

(4) Construct a proof for the right-handed distributive rule, viz. $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

(5) Prove that matrices under addition are commutative and associative.

(6) Prove that $(\mathbf{A}+\mathbf{B})' = \mathbf{A}'+\mathbf{B}'$.

(7) If $\mathbf{A}$ and $\mathbf{B}$ are symmetric, prove that $\mathbf{AB}-\mathbf{BA}$ is skew-symmetric. What can you say about $\mathbf{AB}-\mathbf{BA}$ if $\mathbf{A}$ and $\mathbf{B}$ are both skew-symmetric?

(8) Write the matrix $\begin{pmatrix} 1 & 3 \\ 5 & 2 \end{pmatrix}$ as the sum of a symmetric and a skew-symmetric matrix. Can this always be done for a square matrix?

(9) If $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and the equation $\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0$ has roots 1 and 2 for λ, prove that $\mathbf{A}^2-3\mathbf{A}+2\mathbf{I} = \mathbf{Z}$ where $\mathbf{Z}$ is the 2×2 zero matrix. Deduce that $\mathbf{A}^{-1} = \frac{1}{2}(3\mathbf{I}-\mathbf{A})$.

† Strictly the left-handed distributive rule, as opposed to $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$.

18.9 *Subsets of 2×2 matrices: scalar matrices: C-matrices*

We have already seen that a matrix of the form

$$\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

represents an enlargement. If we take another enlargement matrix

$$\mathbf{B} = \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$$

we see immediately that

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} a+b & 0 \\ 0 & a+b \end{pmatrix} \text{ and } \mathbf{AB} = \begin{pmatrix} ab & 0 \\ 0 & ab \end{pmatrix}.$$

Furthermore, if $a \neq 0$, the matrix

$$\begin{pmatrix} 1/a & 0 \\ 0 & 1/a \end{pmatrix}$$

is the inverse of $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ as the reader will easily verify.

When we compare the properties of real numbers on the one hand with these enlargement matrices (sometimes called scalar matrices) on the other, we see that:

For real numbers	*For scalar matrices*
$(a)+(b) = (a+b)$	$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}+\begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ 0 & a+b \end{pmatrix}$
$a\times 1/a = 1$	$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} 1/a & 0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$(ab)c = a(bc)$	$\begin{pmatrix} ab & 0 \\ 0 & ab \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} bc & 0 \\ 0 & bc \end{pmatrix}$
$ab = ba$	$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$

and we could go on to show that for every property of the real number system there is a corresponding matrix result concerning scalar matrices.

Such a state of affairs, that is, when two algebraic systems have the same structure, is called an *isomorphism*. We say that the set of real numbers is *isomorphic* to the set of scalar matrices $a\mathbf{I}$.

We now try to make capital out of this isomorphism, for a scalar

matrix of the form $\begin{pmatrix}-2 & 0\\ 0 & -2\end{pmatrix}$ involves not only an enlargement, but a rotation through 180° as well. In fact

$$\begin{pmatrix}-2 & 0\\ 0 & -2\end{pmatrix}=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix},$$

so we consider the matrix $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ as typical of a negative scalar matrix. It represents a rotation through 180°, and corresponds to the number -1.

Now, in the real number system, there is no number x with the property that $x^2 = -1$ and there is clearly no member $\mathbf{X}$ of the set of scalar matrices with the property that $\mathbf{X}^2 = -\mathbf{I}$.

But, since $-\mathbf{I}$ represents a rotation of 180°, there are *matrices* (though not scalar ones) which have the property $\mathbf{X}^2 = -\mathbf{I}$. One of these represents a rotation through 90°, viz. $\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, and we see that

$$\begin{pmatrix}0 & -1\\ +1 & 0\end{pmatrix}\begin{pmatrix}0 & -1\\ +1 & 0\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}.$$

We call this matrix $\mathbf{J}$, so that

$$\mathbf{J}^2=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}=-\mathbf{I}.$$

We now consider our two matrices

$$\mathbf{I}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\text{ and }\mathbf{J}=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

and use them as a basis for a vector space of matrices; every matrix of this space will be of the form:

$$a\mathbf{I}+b\mathbf{J}=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}.$$

That this is a useful vector space we now demonstrate; in addition to the usual vector space structure we have a built-in multiplication rule.

For:

$$\begin{aligned}(a\mathbf{I}+b\mathbf{J})(c\mathbf{I}+d\mathbf{J}) &= \begin{pmatrix}a & -b\\ b & a\end{pmatrix}\begin{pmatrix}c & -d\\ d & c\end{pmatrix}\\ &= \begin{pmatrix}ac-bd & -ad-bc\\ bc+ad & ac-bd\end{pmatrix}\\ &= (ac-bd)\mathbf{I}+(ad+bc)\mathbf{J}.\end{aligned}$$

We see that any two matrices of this space have a built-in multiplication rule (the matrix rule), and the product of any two such matrices is also a matrix of this space. Note also it is the same result that we would get if we multiply out as usual, saying $\mathbf{IJ} = \mathbf{JI} = \mathbf{J}$ and $\mathbf{I}^2 = -\mathbf{J}^2 = \mathbf{I}$.

We may also find inverses for

$$\begin{aligned}(a\mathbf{I}+b\mathbf{J})^{-1} = \begin{pmatrix}a & -b\\ b & a\end{pmatrix}^{-1} &= \frac{1}{a^2+b^2}\begin{pmatrix}a & b\\ -b & a\end{pmatrix}\\ &= \frac{a}{a^2+b^2}\mathbf{I}-\frac{b}{a^2+b^2}\mathbf{J}\end{aligned}$$

provided we may divide by a^2+b^2, i.e. a and b are not both zero.

EXERCISE 18h

In this exercise, matrices of the type $\begin{pmatrix}a & -b\\ b & a\end{pmatrix}$ are called **C**-matrices.

(1) Show that there are **C**-matrices with the property $\mathbf{X}^2 = \mathbf{J}$. Interpret them geometrically.

(2) Show that any **C**-matrix can be written in the form **AR**, where **A** is a scalar matrix and **R** is orthogonal. Can it also be written in the form **RA**?

(3) Show that scalar matrices form a commutative group under addition, but only form a group (also commutative) under multiplication if the zero matrix is omitted. Are these properties shared by the real number system?

(4) Establish a left-hand distributive law for scalar matrices, i.e. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$. Why is it unnecessary to establish also a right-hand distributive law? (This completes the isomorphism with the real number system.)

(5) Which of the properties listed in questions 3 and 4 are established also for **C**-matrices? The reader should provide the proofs where necessary.

(6) Having shown that the set of C-matrices, viz. $\{a\mathbf{I}+b\mathbf{J}$: a and b real numbers$\}$, has itself all the properties of a number system, it is left to the reader to show that it is in fact isomorphic with the system of complex numbers $a+bj$.†

(7) The set of all square matrices of order n is mapped into the real numbers under the mapping $\mathbf{M} \to \det \mathbf{M}$. Discuss how the images of $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_1\mathbf{M}_2$ and $\mathbf{M}_2\mathbf{M}_1$ are related. Is there any connection to be seen between failure to 'divide by zero' among real numbers on one hand and any similar type of failure among the matrices? What subset of the matrices maps into the number zero?

(8) Set up tables for addition and for multiplication of the set of integers $\{0, 1, 2, 3, 4\}$ modulo 5. Are both these tables commutative group tables? Prove any further property which is needed to establish this system as a field (see ref. 11 on p. 213).

18.10 *The algebra with matrices as elements*

In chapter 1 we carefully stated the minimum conditions for an algebra, and followed this up by laying down the foundations for the algebra of vectors. We now have the algebra of matrices, themselves a kind of super-vector, which have the property of being an operating mechanism on vectors. But this operation is important in application: the pure algebraist would say that he has another system, richer in properties than before.

In section 6.6 we defined the algebraic system known as the vector space. We have already seen that all the definitions are satisfied when the elements are matrices (of a specific size). However, the vector space structure only deals with addition of matrices and multiplication by a scalar, and gives no account of matrix multiplication.‡

The structure to which square matrices belong is called a linear algebra. For a linear algebra the elements have to form a vector space, with a multiplication defined which is distributive and associative. Clearly square matrices satisfy these conditions.

Particular subsets of square matrices may have a rather stronger algebraic structure. For example, if we restrict attention to non-singular matrices we see that the conditions for a group under multiplication are satisfied, and the reader will see in the examples which follow that if we restrict even further we will find further 'pockets' of algebraic structure.

† The algebraic term for a system with such properties is a *field*.

‡ We have seen, however, that there is nothing to prevent a further law of combination existing, as happened for the C-matrices in section 18.9.

EXERCISE 18i (*Miscellaneous examples on matrices*)

(1) Prove that, when n is a positive integer,

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}.$$

Interpret this result in the complex number system.

(2) Write down 2×2 matrices which represent a general rotation, and a general reflection. Calling these matrices $\mathbf{X}$ and $\mathbf{Y}$ respectively, show that $\mathbf{XY}$ and $\mathbf{YX}$ are both reflections.

Hence show geometrically that 2×2 orthogonal matrices form a group under multiplication.

(3) Show that if $\mathbf{A}^2 = \mathbf{A}$, and $\mathbf{A} \neq \mathbf{I}$, then $\mathbf{A}$ is singular. Is the same true if $\mathbf{A}^3 = \mathbf{A}$? Give examples to illustrate your answer.

(4) We have already seen that matrices of a particular size form a vector space. If we take the set of 2×2 matrices, what would be the dimension (defined in chapter 6) of this space? Suggest a basis.

(5) Prove that $$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

Interpret this result geometrically. Referring to exercise 18e, question 7, which type of matrix are we dealing with?

(6) A matrix $\mathbf{E}$ such that $\mathbf{E}^2 = \mathbf{E}$ is called a 'projection'. Give an example of such a matrix and show that $\mathbf{E}$ is a projection $\Leftrightarrow \mathbf{I}-\mathbf{E}$ is a projection.

(7) Look up the definition of equivalence.

If $\mathbf{A}$ and $\mathbf{B}$ are matrices of the same size $(m\times n)$ and non-singular matrices $\mathbf{P}(m\times m)$ and $\mathbf{Q}(n\times n)$ can be found such that $\mathbf{B} = \mathbf{PAQ}$, show that non-singular matrices $\mathbf{R}$ and $\mathbf{S}$ can also be found so that $\mathbf{A} = \mathbf{RBS}$.

Prove further that in the set of $m\times n$ matrices, those which can be related to $\mathbf{A}$ and $\mathbf{B}$ in this way form an equivalence class.

(8) Show that the inverse of a matrix may be found by performing elementary column operations instead of elementary row operations, but not by mixing the two. Why not?

(9) Show that after a square matrix has been reduced to echelon form using row operations, it may always be further reduced using column operations to a form where the only non-zero elements are on the main diagonal, and these are unity. Explain why, with an example, column operations may be necessary at all. If this final matrix has r zeros on the main diagonal and is called $\mathbf{I}_r$, what is the matrix relationship

between it and the original matrix **A**? (*Hint*: Remember that an elementary operation has a corresponding elementary matrix.) Show that this is an equivalence relation.

(10) Before performing a matrix mapping **T** on a vector space, it is often convenient first to change the basis so that the matrix mapping now required (**S**) is simpler: a method of changing the basis is here considered. If the basis vectors of a three-space were originally

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} \text{ and } \begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\end{pmatrix} \text{ and the new basis is } \begin{pmatrix}a_1\\ a_2\\ a_3\end{pmatrix}\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix} \text{ and } \begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix},$$

show that *if*

$$\begin{pmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix} = \mathbf{M}\begin{pmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{pmatrix}$$

M is non-singular.

If any transformation matrix is **T**, and the matrix performing the same transformation but relative to the new basis is **S**, prove that $\mathbf{T} = \mathbf{M}^{-1}\mathbf{S}\mathbf{M}$.

Show that matrices of the same size related as are **S** and **T** above form an equivalence class.

(11) The quadratic equation $ax^2+bx+c = 0$ may be written in the form

$$(x \quad 1)\begin{pmatrix}a & b/2\\ b/2 & c\end{pmatrix}\begin{pmatrix}x\\ 1\end{pmatrix} = (0).$$

What is the significance of the vanishing of the determinant?

(12) Prove that, if a matrix **P** exists such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{I}_2$ where

$$\mathbf{I}_2 = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

then $\mathbf{A}^2 = \mathbf{A}$.

(13) Investigate the matrix equations $\mathbf{AX} = \mathbf{B}$ in each of the following cases:

(*i*) $\mathbf{A} = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$, $\mathbf{B} = \begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix}$

(*ii*) $\mathbf{A} = \begin{pmatrix}2 & 1\\ 4 & 2\end{pmatrix}$, $\mathbf{B} = \begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix}$

(*iii*) $\mathbf{A} = \begin{pmatrix}2 & 1\\ 4 & 2\end{pmatrix}$, $\mathbf{B} = \begin{pmatrix}3 & 6\\ 6 & 12\end{pmatrix}$.

(14) If a, b, c and d are all positive, prove that there is a positive value of t such that the equations

$$ax+by = tx,$$
$$cx+dy = ty$$

have solutions other than $x = y = 0$, and that there are solutions corresponding to this value of t in which both x and y are positive.

(*Cambridge Scholarship*)

Harder Examples on Matrices

(1) Consider the matrices

$$\mathbf{X} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix}\mathbf{P} & \mathbf{Q}\\ \mathbf{R} & \mathbf{S}\end{pmatrix}$$

where the elements of **X** and **Y** are themselves the matrices

$$\mathbf{A} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\ \mathbf{B} = \begin{pmatrix}2 & 3\\ 1 & 1\end{pmatrix},\ \mathbf{C} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\ \mathbf{D} = \begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix}$$
$$\mathbf{P} = \begin{pmatrix}2 & 0\\ 4 & 3\end{pmatrix},\ \mathbf{Q} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\ \mathbf{R} = \begin{pmatrix}2 & 6\\ 3 & 1\end{pmatrix},\ \mathbf{S} = \begin{pmatrix}4 & 0\\ 1 & -1\end{pmatrix}.$$

Test whether the usual rule for matrix multiplication works, i.e. whether

$$\mathbf{XY} = \begin{pmatrix}\mathbf{AP}+\mathbf{BR} & \mathbf{AQ}+\mathbf{BS}\\ \mathbf{CP}+\mathbf{DR} & \mathbf{CQ}+\mathbf{DS}\end{pmatrix}.$$

(2) Consider the matrix

$$\mathbf{A} = \begin{pmatrix}3 & 3\\ 2 & 4\end{pmatrix}$$

where the numbers are to be combined according to the rules of arithmetic modulo 5. Show that the matrix is non-singular, and find its inverse. Consider the same problem modulo 6.

(3) It is given that the matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & -3 & 1 \\ 3 & -1 & 4 & -3 \\ 5 & -4 & 11 & -7 \\ 7 & 0 & 5 & -5 \end{pmatrix}$$

is singular. Calculate the values of each of the cofactors of the bottom row.

Reduce the matrix to echelon form. Is there any relation between the two results?

(4) If $\mathbf{A}$ is *any* 2×2 matrix show that any matrix which commutes with $\mathbf{A}$ is of the form $\lambda\mathbf{A}+\mu\mathbf{I}$. For the particular case

$$\mathbf{A} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix},$$

write $\mathbf{A}^2$ and $\mathbf{A}^{-1}$ in this form.

(5) The (1-1) algebraic correspondence

$$att'+bt+ct'+d = 0$$

may be written $\dfrac{t}{1} = \dfrac{-ct'-d}{at'+b}$.

If we write this as

$$\begin{pmatrix} t \\ 1 \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}\begin{pmatrix} t' \\ 1 \end{pmatrix},$$

does the 'true matrix inverse' of $\begin{pmatrix} -c & -d \\ a & b \end{pmatrix}$ give us the inverse correspondence

$$\begin{pmatrix} t' \\ 1 \end{pmatrix} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}\begin{pmatrix} t \\ 1 \end{pmatrix}?$$

If we had another correspondence $\begin{pmatrix} T \\ 1 \end{pmatrix} = \begin{pmatrix} -C & -D \\ A & B \end{pmatrix}\begin{pmatrix} t \\ 1 \end{pmatrix}$, would the usual matrix product work? What is the significance of the vanishing of the determinant?

(6) For the central conic (i.e. conic with centre at the origin) $ax^2+2hxy+by^2 = 1$, show that the gradient of the normal at any point (X, Y) is $\dfrac{hX+bY}{aX+bY}$. If the normal at this point passes through the origin show that

$$\begin{pmatrix} a & h \\ h & b \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} = \lambda\begin{pmatrix} X \\ Y \end{pmatrix}$$

where λ is a scalar.

Show how this line of thought may be developed to find the directions of the axes of the conic $41x^2-24xy+34y^2 = 1$.

(7) A drunkard is at the 'Bull' public house which is $1\frac{1}{2}$ miles from home. He starts to go home, but walks only $\frac{1}{2}$ mile and then wonders where he is. His chance of going on for another $\frac{1}{2}$ mile is the same as his chance of going back to the Bull. The same thing happens after each half mile walked, with the proviso that if he gets home he stays there, and if he gets back to the Bull he stays there.

We can set up a vector corresponding to the probabilities of his being in any one of his stopping-places as follows:

Probability of being at Bull	p_1	
Probability of being $\frac{1}{2}$ mile from Bull	p_2	where
Probability of being $\frac{1}{2}$ mile from home	p_3	$p_1+p_2+p_3+p_4 = 1$.
Probability of being home	p_4	

(*a*) What is the corresponding vector at the next stage?

(*b*) Write down the matrix, $\mathbf{M}$, which transforms the first vector into the second, whatever the values of p_1, p_2, p_3 and p_4.

(*c*) Calculate $\mathbf{M}^2$, $\mathbf{M}^4$ and $\mathbf{M}^8$ and verify that

$$\mathbf{M}^8 = \begin{pmatrix} 1 & 85/128 & 85/256 & 0 \\ 0 & 1/256 & 0 & 0 \\ 0 & 0 & 1/256 & 0 \\ 0 & 85/256 & 85/128 & 1 \end{pmatrix}.$$

(*d*) What is the probability vector $\mathbf{p}$ just before he makes his first decision?

(*e*) Evaluate $\mathbf{M}^8\mathbf{p}$ and so find out the probabilities of his being in the various places after 8 decisions.

(*f*) What do you think are his ultimate chances of getting home?

(*g*) How is the question altered if the landlord of the Bull sends him home again?

(8) If ϕ is a function of x and y, and $x = r\cos\theta$ and $y = r\sin\theta$ so that ϕ may also be considered as a function of r and θ, then we know that

$$\frac{\partial\phi}{\partial\theta} = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial\theta}$$

and
$$\frac{\partial\phi}{\partial r} = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial\phi}{\partial y}\frac{\partial y}{\partial r}$$

so that we may write, using the usual notation,

$$\begin{pmatrix}\phi_\theta\\ \phi_r\end{pmatrix} = \begin{pmatrix}x_\theta & y_\theta\\ x_r & y_r\end{pmatrix}\begin{pmatrix}\phi_x\\ \phi_y\end{pmatrix}.$$

What is the value of the determinant of this 2×2 matrix in this instance?

(9) If z is a function of x and y, the total differential is given by

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy.$$

This is reminiscent of the scalar product of two vectors

$$\mathbf{i}\frac{\partial z}{\partial x} + \mathbf{j}\frac{\partial z}{\partial y} \quad\text{and}\quad dx\,\mathbf{i} + dy\,\mathbf{j}.$$

The first of these vectors is called grad **z** and the second **dr**. Then $dz = \text{grad } \mathbf{z} \,.\, \mathbf{dr}$.

For the surface $z = 3x^2 + 2y^2 - 2x$, find:

(*i*) grad **z** at the point (2, 1, 10)
(*ii*) an expression for dz at (2, 1, 10)
(*iii*) the equation of the tangent plane at that point.

(10) Find the inverse of $\mathbf{A} = \begin{pmatrix}1 & 2 & 7\\ 1 & 3 & 0\\ 0 & -1 & 8\end{pmatrix}$. Use your answer to invert $\mathbf{B} = \begin{pmatrix}24 & -8 & -1\\ -23 & 8 & 1\\ -21 & 7 & 1\end{pmatrix}$, proving any general result you use.

(11) By trying as solutions $x = Ae^{\lambda t}$, $y = Be^{\lambda t}$, find permissible values of λ and the corresponding ratios $A{:}B$, for the simultaneous equations

$$\frac{dx}{dt} = 2x + 4y$$
$$\frac{dy}{dt} = 3x + y.$$

Show that the general solution is

$$x = ce^{-2t} + 4de^{5t}$$
$$y = -ce^{-2t} + 3de^{5t}$$

and find the particular solution such that $x = 4$ when $t = 0$ and $y \to 0$ as $t \to \infty$.

(12) Show that the 3-vectors which are mapped into the zero vector by the matrix

$$\mathbf{M} = \begin{pmatrix}3 & 2 & -1\\ 2 & -4 & 3\\ 1 & -10 & 7\end{pmatrix}$$

form a vector space in their own right. Find a basis for this space.

(13) Show that all functions of t which form solutions of the differential equation

$$\frac{d^2x}{dt^2} - 3\frac{dx}{dt} + 2x = 0$$

form a vector space, and state a basis for this space.

(14) Show that all 2×2 matrices form a vector space of dimension 4 and suggest a simple form of basis for the space. An important sub-space of dimension 2 exists with **I** and **J**, $= \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, as basis: show that if the matrix **B** commutes with any two distinct members of this sub-space then (*i*) it does so with all, and (*ii*) it is itself a member of this sub-space.

(15) Show that any 2×2 matrix **M** can be written in the form $\begin{pmatrix}a+b & c-d\\ c+d & a-b\end{pmatrix}$, and hence express it as a member of a space of dimension 4 with **I** and three other orthogonal matrices **X**, **Y**, **J** as basis. Show that (*i*) $\mathbf{X}^2 = \mathbf{Y}^2 = -\mathbf{J}^2 = \mathbf{I}$, (*ii*) $\mathbf{RS} = -\mathbf{SR}$ for any pair of **X**, **Y**, **J**.

Taking the special case in which $a^2 - b^2 + c^2 - d^2 = 1$, show that $\mathbf{M}^{-1}$ is obtained by changing the signs of b, c, and d.

(16) By considering $\sum_{i=1}^{n}(a_i-\lambda b_i)^2$ as a function of λ, or otherwise, prove that $\sum_{i=1}^{n}a_i^2\sum_{i=1}^{n}b_i^2 \geqslant (\sum_{i=1}^{n}a_ib_i)^2$. Hence show that $|\mathbf{a}+\mathbf{b}| \leqslant |\mathbf{a}|+|\mathbf{b}|$. Interpret this result geometrically in space of 2 and 3 dimensions.

(17) Show that the simultaneous equations

$$\begin{aligned} ax+\ y+\ z &= p,\\ x+ay+\ z &= q,\\ x+\ y+az &= r \end{aligned}$$

have a unique solution if a has neither the values -1 or 2. Show also that, if $a=-2$, there is no solution unless p, q and r satisfy a certain condition, and that there are then an infinite number of solutions. Discuss the solution of the equations when $a=1$.

Find the most general solution (if any) in the following cases: (*i*) $a=3$, $p=q=r=1$, (*ii*) $a=-2$, $p=q=r=1$, (*iii*) $a=-2$, $p=1$, $q=-1$, $r=0$, (*iv*) $a=1$, $p=q=r=0$.

(*Cambridge Scholarship*)

(18) Show that, if $\lambda=3$, it is possible to choose constants α, β and γ, not all zero, such that

$$\alpha(11x-6y+2z)+\beta(-6x+10y-4z)+\gamma(2x-4y+6z)$$

is identically equal to

$$\lambda(\alpha x+\beta y+\gamma z).$$

Obtain the ratios of α, β and γ. Find all the other values of λ for which it is possible to find constants α, β and γ, not all zero, with the above property.

(*Cambridge Scholarship*)

(19) $f(x)$, $g(x)$ and $h(x)$ are functions of x satisfying the equations

$$\frac{df}{dx}=f+g+2h,$$

$$\frac{dg}{dx}=2g+2h,$$

$$8\frac{dh}{dx}=7f+8g+24h.$$

Show that $f-g$ is of the form $Ae^{\lambda x}$, where A is a constant.

Find all the solutions of the form

$$f=f_0e^{\lambda x},\ g=g_0e^{\lambda x},\ h=h_0e^{\lambda x},$$

where λ is independent of x, giving the possible values of λ and the corresponding ratios of the constants f_0, g_0, h_0.

(*Cambridge Scholarship*)

(20) In the system of equations

$$\begin{aligned} -ny+mz &= a,\\ nx\quad\ -lz &= b,\\ -mx+ly\quad &= c,\\ lx+my+nz &= p, \end{aligned}$$

l, m, n, a, b, c, p are given real numbers, and l, m, n are not all zero. Prove that a necessary and sufficient condition for the equations to have a solution is that

$$la+mb+nc=0,$$

and solve the equations when this condition is satisfied.

(*Cambridge Scholarship*)

(21) The nine numbers a_{ij} $(i,j=1,2,3)$ satisfy the equations

$$a_{i1}a_{j1}+a_{i2}a_{j2}+a_{i3}a_{j3}=\delta_{ij}\quad (i,j=1,2,3),$$

where $\delta_{ij}=0$ if $i\neq j$ but $\delta_{ij}=1$ if $i=j$. Show that they also satisfy the equations

$$a_{1i}a_{1j}+a_{2i}a_{2j}+a_{3i}a_{3j}=\delta_{ij}\quad (i,j=1,2,3).$$

Prove also that $a_{22}a_{33}-a_{23}a_{32}=\pm a_{11}$.

(*Cambridge Scholarship*)

(22) Given three real non-zero numbers a, b and h, prove that the relations

$$\begin{aligned} ax+hy &= \lambda x,\\ hx+by &= \lambda y \end{aligned}$$

can be satisfied by two distinct real values of λ, and for each of these values of λ there exists a definite value of the ratio x/y.

By considering $\lambda_1^2+\lambda_2^2$, or otherwise, where λ_1 and λ_2 are the two values of λ, prove that the numerical values of λ_1 and λ_2 cannot exceed $\sqrt{(a^2+b^2+2h^2)}$.

Is it possible for λ_1 or λ_2 to have this extreme value?

(*Cambridge Scholarship*)

(23) If

$$x_1x_2+y_1y_2 = a_1,\ x_2x_3+y_2y_3 = a_2,\ x_3x_1+y_3y_1 = a_3,$$
$$x_1^2+y_1^2 = x_2^2+y_2^2 = x_3^2+y_3^2 = b,$$

prove that

$$b^3-(a_1^2+a_2^2+a_3^2)b+2a_1a_2a_3 = 0.$$

Deduce that, if the $2n$ equations

$$x_1x_2+y_1y_2 = a_1,\ x_2x_3+y_2y_3 = a_2.\ \ldots$$
$$\ldots\ x_{n-1}x_n+y_{n-1}y_n = a_{n-1},\ x_nx_1+y_ny_1 = a_n,$$
$$\text{and } x_1^2+y_1^2 = x_2^2+y_2^2 = \ldots = x_n^2+y_n^2 = b$$

for the $2n$ unknowns $x_1, \ldots x_n, y_1 \ldots y_n$ are consistent, there must be an algebraic relation connecting b and $a_1, a_2, \ldots a_n$.

(*Cambridge Scholarship*)

APPENDIX

The Multiplication Rule for Matrices

We start from the definition of a matrix as a way of conveniently representing a linear mapping of a vector space.

Suppose that a linear mapping operating on the vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ maps it into $\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$. We know that a relationship exists of the type

$$\xi = a_1x+b_1y+c_1z$$
$$\eta = a_2x+b_2y+c_2z$$
$$\zeta = a_3x+b_3y+c_3z.$$

We are entitled to write the coefficients in a unified pattern $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$, to denote this whole pattern by a single letter $\mathbf{M}$ if desired, and to denote the operation of mapping by $\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

In full, this would be

$$\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and the right-hand side is therefore *by definition* equivalent to the column vector

$$\begin{pmatrix} a_1x+b_1y+c_1z \\ a_2x+b_2y+c_2z \\ a_3x+b_3y+c_3z \end{pmatrix}.$$

Similarly,

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{pmatrix}\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \text{ or } \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{N}\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

We see immediately from the matrix representation that

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{N}\left\{\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right\} = (\mathbf{NM})\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

and it remains to find out what the relation between $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ and $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is in its linear equation form.

Substituting from the linear equations, we have

$$\begin{aligned} X &= p_1\xi+q_1\eta+r_1\xi \\ &= p_1(a_1x+b_1y+c_1z)+q_1(a_2x+b_2y+c_2z)+r_1(a_3x+b_3y+c_3z) \\ &= (p_1a_1+q_1a_2+r_1a_3)x+(p_1b_1+q_1b_2+r_1b_3)y+(p_1c_1+q_1c_2+r_1c_3)z \end{aligned}$$

with similar expressions for Y and Z.

Re-writing in matrix notation, we have

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} p_1a_1+q_1a_2+r_1a_3 & p_1b_1+q_1b_2+r_1b_3 & p_1c_1+q_1c_2+r_1c_3 \\ p_2a_1+q_2a_2+r_2a_3 & p_2b_1+q_2b_2+r_2b_3 & p_2c_1+q_2c_2+r_2c_3 \\ p_3a_1+q_3a_2+r_3a_3 & p_3b_1+q_3b_2+r_3b_3 & p_3c_1+q_3c_2+r_3c_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

But this is the same relation as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = (\mathbf{NM})\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Hence

$$\mathbf{NM} = \begin{pmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{pmatrix}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

$$= \begin{pmatrix} p_1a_1+q_1a_2+r_1a_3 & p_1b_1+q_1b_2+r_1b_3 & p_1c_1+q_1c_2+r_1c_3 \\ p_2a_1+q_2a_2+r_2a_3 & p_2b_1+q_2b_2+r_2b_3 & p_2c_1+q_2c_2+r_2c_3 \\ p_3a_1+q_3a_2+r_3a_3 & p_3b_1+q_3b_2+r_3b_3 & p_3c_1+q_3c_2+r_3c_3 \end{pmatrix},$$

and the product rule is proved in this instance.

We may also deduce how to multiply a matrix by a scalar. For, if $\mathbf{M}$ is the matrix used previously, $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and k is a scalar multiplier, $\mathbf{M}(k\mathbf{v}) = k(\mathbf{Mv})$, as we have already proved that a matrix mapping is linear, and it is now natural to write this as $(k\mathbf{M})\mathbf{v}$, defining the matrix $k\mathbf{M}$. But when we multiply out:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} kx \\ ky \\ kz \end{pmatrix} = \begin{pmatrix} a_1kx+b_1ky+c_1kz \\ a_2kx+b_2ky+c_2kz \\ a_3kx+b_3ky+c_3kz \end{pmatrix}$$

$$= \begin{pmatrix} a_1k+b_1k+c_1k \\ a_2k+b_2k+c_2k \\ a_3k+b_3k+c_3k \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

$$\text{Hence } k\mathbf{M} = k\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} ka_1 & kb_1 & kc_1 \\ ka_2 & kb_2 & kc_2 \\ ka_3 & kb_3 & kc_3 \end{pmatrix}.$$

ANSWERS

Exercise 1a

(*1*) Yes, all the requirements are satisfied.
(*2*) (*i*) Yes (*ii*) In general, no (*iii*) Yes.
(*3*) (*i*) No, no closure. (*ii*) Yes (*iii*) Yes (*iv*) (*a*) Yes (*b*) No, not closed.
(*4*) (*i*) No, not defined for all possible pairs of points. (*ii*) No. Still fails if AB $\parallel$ l (*iii*) Yes (*iv*) Yes.
We can test *all* cases, even (*i*) and (*ii*) when they work, to see whether they are commutative (C) and associative (A). Results are:
(*i*) C, A (*ii*) C, A (*iii*) C, not A (*iv*) Not C, not A.
(*5*) (*ii*) They are not the same. (*iii*) $A n \varnothing$ and $A n A$ are both the set of all points in $\mathscr{E}$ and not in A, i.e. $(\mathscr{E} - A)$. $A n \mathscr{E}$ is $\varnothing$.
(*6*) (*i*) A (*ii*) $\varnothing$ (*iii*) A, A, $\mathscr{E}$ (*iv*) A, A, $\varnothing$ (*v*) $\varnothing$ (*vi*) A − B.
(*7*) $p \vee q$ is (*a*). $p \wedge q$ is (*c*). not-$p \wedge$ not-q is (*b*). $(p \wedge q) \vee r$ is (*c*).

Exercise 1b

(*1*) For all A, $A \cup \varnothing = A$ and $\varnothing \cup A = A$, $A - \varnothing = A$ but $\varnothing - A \neq A$.
(*2*) Unity; but not for division, since $1 \div n \neq n$.
(*3*) $A n B$ cannot ever be equal to A or to B since it contains elements which are neither in A nor in B.
(*4*) It is true that $A \,\Delta\, (B \,\Delta\, C) = (A \,\Delta\, B) \,\Delta\, C$, associativity.

Exercise 1c

(*1*) Zero (*i*) 8 (*ii*) 6, which is the 'opposite' of 4; yes.
(*2*) 6; zero has no inverse.
(*3*) Unity. Inverse is reciprocal.
(*4*) (*a*) $I = \varnothing$ (*b*) $Z = Y$, i.e. each set is its own inverse.
(*6*) Unity is the identity: 1 and 9 are inverses of themselves, 3 and 7 of each other. Associative, e.g. $(3\times7)\times9 = 3\times(7\times9)$.
(*8*)

	Closure	*Neutral*	*Inverses*	*Commute*	*Associate*
(*a*)	✓	✓(−2)	✓	✓	✓
(*b*)	✓	✓(0)	✓(self)	✓	×
(*c*)	×	×	.	✓	×
(*d*)	✓	✓(0)	×	✓	×

(*9*) $(ac)d \neq a(cd)$; $(ab)d \neq a(bd)$
$\{(ab)c\}d = a$, $(ab)(cd) = a$, $\{a(bc)\}d = d$, $a\{(bc)d\} = b$.

$\langle1\rangle\ \langle2\rangle\ \langle3\rangle\ \langle0\rangle$
(*10*) The table: $\langle2\rangle\ \langle3\rangle\ \langle0\rangle\ \langle1\rangle$; and the expressions $\langle1\rangle$, $\langle3\rangle$.
$\langle3\rangle\ \langle0\rangle\ \langle1\rangle\ \langle2\rangle$
$\langle0\rangle\ \langle1\rangle\ \langle2\rangle\ \langle3\rangle$

Exercise 2a

(*1*) (*a*) 096°, 37·2 miles; (*b*) 37·0 miles east, 4·0 miles south.
(*2*) N 37° W (or 323°), $2\frac{1}{2}$ miles.
(*3*) N 49° E (or 049°), $\sqrt{7} \simeq 2{\cdot}6$ m.p.h.
(*4*) (*a*) $056\frac{1}{2}$°, 86·5 miles; (*b*) (72·1, 47·8).
(*5*) 490 knots, 306°
(*6*) 184°, 464 knots
} tolerances if drawn: ± 5 knots, ± 1 degree.

Exercise 2b

(*1*) Diagonals of a parallelogram . . . .
(*2*) $\begin{pmatrix}0\\0\\0\end{pmatrix}$; $\mathbf{s} = \begin{pmatrix}-2{\cdot}7\\2{\cdot}1\\3{\cdot}7\end{pmatrix}$; $\mathbf{s}' = \begin{pmatrix}2{\cdot}7\\-2{\cdot}1\\-3{\cdot}7\end{pmatrix}$; $(\mathbf{a}+\mathbf{b})+\mathbf{s}' = \mathbf{a}+(\mathbf{b}+\mathbf{s}') = \mathbf{0}$.
(*4*) Addition is not by components, e.g. £1 18s. 10d. + £1 5s. 3d. has the sum £3 4s. 1d, not £2 23s. 13d.
Also, only a limited range of values of 'components' is here permissible.

Exercise 3a

(2) $\mathbf{b}-\mathbf{a}$, $\mathbf{b}+(-\mathbf{a})$, $-\mathbf{a}+\mathbf{b}$ (*3*) $\frac{1}{3}(-\mathbf{a}+\mathbf{b}) = \frac{1}{3}(-\mathbf{a})+\frac{1}{3}\mathbf{b}$; 1 to 3.
(*4*) Midpoint of *AB*, *M* say; midpoint of *MC*; point of trisection of *MC* nearer *M*, i.e. centroid of $\triangle ABC$.
(*5*) (*i*) $-\mathbf{a}+\mathbf{c}$, $-2\mathbf{a}$, $-2\mathbf{a}-\mathbf{c}$; (*ii*) $2\mathbf{a}+2\mathbf{c}$, $-2\mathbf{c}$, $-\mathbf{a}+\mathbf{c}$.

Exercise 3b

(2) (*i*) 1, 1 (*ii*) 6, $\frac{1}{2}$ or -6, $-\frac{1}{2}$ (*iii*) zero, $\frac{2}{3}$ (*iv*) $-\frac{1}{2}$, 2 (*v*) 2.
(*3*) $\mathbf{p}-\mathbf{q} = 2\mathbf{a}+(-2)\mathbf{b}$; $\mathbf{q}-\mathbf{p} = (-2)\mathbf{a}+2\mathbf{b}$; $\mathbf{p}+\frac{1}{2}\mathbf{q} = 5\mathbf{a}+\frac{5}{2}\mathbf{b}$;
$\frac{1}{2}(\mathbf{p}-\mathbf{q}) = 1.\mathbf{a}+(-1)\mathbf{b}$; $\mathbf{a} = \frac{3}{10}\mathbf{p}+(-\frac{1}{10})\mathbf{q}$; $\mathbf{b} = (-\frac{1}{5})\mathbf{p}+\frac{2}{5}\mathbf{q}$.
(*4*) (*a*) $\mathbf{a} = \begin{pmatrix}2\\1\end{pmatrix}$, (*b*) $\mathbf{b} = \begin{pmatrix}2\\-3\end{pmatrix}$; (*c*) $\overline{AD} = \mathbf{b} = \overline{BC}$; $\overline{DC} = \mathbf{a}$; $\overline{BA} = -\mathbf{a}$,
$\overline{CB} = -\mathbf{b}$, $\overline{AC} = \mathbf{a}+\mathbf{b}$, $\overline{CA} = -\mathbf{a}-\mathbf{b}$; $\overline{AE} = \frac{1}{2}\mathbf{a}+\frac{1}{2}\mathbf{b} = \overline{EC}$,
$\overline{BE} = -\frac{1}{2}\mathbf{a}+\frac{1}{2}\mathbf{b} = \overline{ED}$.

Exercise 3c (Answers to 1 decimal place and nearest degree)

(*1*) As 3-vectors, components to 1 decimal place, 3rd direction being upwards:
(*i*) $\mathbf{u} = \begin{pmatrix}26{\cdot}0\\15{\cdot}0\\0\end{pmatrix}$; (*ii*) $\mathbf{v} = \begin{pmatrix}0\\-10{\cdot}0\\1{\cdot}0\end{pmatrix}$; (*iii*) $\mathbf{u}+\mathbf{v} = \begin{pmatrix}26{\cdot}0\\5{\cdot}0\\1{\cdot}0\end{pmatrix}$ of magnitude $\sqrt{702} \simeq 26{\cdot}5$ knots.
(*2*) 25 m.p.h.; 18·0 m.p.h. from N 56° E (or 056°); 48° on each side of south, i.e. 132° to 228°.
(*3*) Like the wind in question 2, from slightly ahead.
(*4*) (*i*) 130 ft; (*ii*) $\begin{pmatrix}0\\50\\120\end{pmatrix}\begin{pmatrix}40\\-30\\120\end{pmatrix}\begin{pmatrix}-40\\-30\\120\end{pmatrix}$; (*iii*) $\begin{pmatrix}0\\500\\1200\end{pmatrix}\begin{pmatrix}333\\-250\\1000\end{pmatrix}\begin{pmatrix}-333\\-250\\1000\end{pmatrix}$.
(*5*) Allowing for the butter and sugar, we have to make up the vector (90, 45, 230) from meat, bread, eggs. This requires almost all the bread, with say 60 of protein from meat and eggs. For vector (x, y, z) the calorie value is

$5x+10y+\frac{9}{2}z$. In practice, some of the sugar in the above diet would be replaced by starch foods.
(*6*) $\mathbf{d} = \frac{3}{2}\mathbf{a}+\frac{1}{2}\mathbf{b}$; $\frac{1}{2}(\mathbf{c}+\mathbf{d}) = \frac{5}{4}\mathbf{a}+\frac{3}{4}\mathbf{b}$; 1 to 2.
(*7*) $\dfrac{\sin 100°}{\sin 20°} \times 30$ m.p.h. $= 86$ m.p.h.

Exercise 4a

(*1*) (*a*) 5 (*b*) 54 (*c*) 34 (*d*) 0. (*2*) $\frac{2}{3}$, $\frac{1}{3}$.
(*3*) (*a*) 5 (*b*) 54 (*c*) 34 (*d*) 0.
(*4*) (*ii*) $\lambda:\mu = 2:1$; (*iv*) $\lambda = 0$, μ any value; (*i*) and (*iii*) independent.
(*7*) -7, $-7k$; (*a*) 0 (*b*) 288 (*c*) 91.
(*8*) (*a*), (*f*) have unique solutions; (*c*) inconsistent.
(*9*) -6, 15. (*10*) $\lambda = +1$, $\mu = 5$; $\lambda = -1$, $\mu = -5$.
(*11*) (*i*) and (*ii*) $\Rightarrow (x+\lambda)\mathbf{a}+(y+\mu)\mathbf{b} = \mathbf{c}$; $x'\mathbf{a}+y'\mathbf{b} = \mathbf{c}$ where x', y' is not the same pair of values as x, y; (*ii*) is false.
(*12*) Valid: $d \Rightarrow$ not-s; $s \Rightarrow$ not-d.
(*13*) p is d; q is not-s OR p is s; q is not-d.

Exercise 4b

(*1*) -7, 5 (*2*) 9, 2 (*3*) -1, 5 (*4*) 4, -3.

Exercise 4c

(*1*) $\simeq 0{\cdot}34$, $0{\cdot}62$ (*2*) $\simeq 0{\cdot}9$, $-0{\cdot}1$ (*3*) 0·3366, 0·6236; 0·9024, $-0{\cdot}09995$.

Exercise 5a

(*1*) (*i*) -2, 1, 4 (*ii*) 3, 0, 1 (*3*) -4, -1, $5\frac{1}{2}$.

Exercise 5b

(*1*) 3, -2, 0. $\Delta = 83$ if final step is $r_3' = r_3+\frac{8}{3}r_2$; but if $r_3' = 3r_3+8r_2$ we get $\Delta = 249$.
In questions (*4*) to (*9*) the echelon forms are not unique, but all are satisfied by the following solutions:
(*4*) 1, -1, 1 (*5*) 1, -2, 1 (*6*) -1, -5, 1
(*7*) -2, 1, 1 (*8*) 3, 2, 1 (*9*) -2, 1, -1

Exercise 6a

(*1*) * (*a*) -1 (*b*) $-2-a$ (*c*) Yes (*d*) Yes.
° (*a*) 0 (*b*) No (*c*) Yes (*d*) No.
(*2*) Yes; 0, 5, 4, 3, 2, 1.
(*3*) No; 1 is identity but 0, 2, 3, 4 have no inverses.

Exercise 6b

(*1*) $-3:1:1$. (*3*) 3, 3 (Select the second vector and any two others).

Exercise 6d

(*1*) (*a*) No $-$ve x, 1-many. (*b*) No $|x|$, $|y| > 1$, many-many. (*c*) All values, many-1 *and* 1-many (could be called many-many). (*d*) All values, many-many. (*e*) All values, 1-1.

(*2*) (*a*) Only 0, ± 3, ± 4, ± 5; many-many. (*b*) Only $p = \pm 1$, $q = 0$, many-1. (*c*) Odd p only, 1-1. (*d*) p, all values; q only ± 1 and 0; many-1. (*e*) All values, many-many ($p = q \pm 1$). (*f*) $p = q = 0$ only; 1-1.

(*3*) (*a*) (*b*) (*c*) All values, 1-1. (*d*) All values of **a**; only those of α with modulus 0 or 1; many-1.

Exercise 7a

(*1*) (*a*) 5 (*b*) 5 (*c*) 5 (*d*) 5 (*e*) 13 (*f*) 13 (*g*) 1 (*h*) 3 (*i*) $\sqrt{3}$ (*j*) 1.

(*2*) 2 edges given by each of $\begin{pmatrix} 2a \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 2a \\ 0 \end{pmatrix}\begin{pmatrix} a \\ a \\ \sqrt{2}a \end{pmatrix}\begin{pmatrix} -a \\ a \\ \sqrt{2}a \end{pmatrix}\begin{pmatrix} a \\ -a \\ \sqrt{2}a \end{pmatrix}\begin{pmatrix} -a \\ -a \\ \sqrt{2}a \end{pmatrix}$

or those with reversed signs. Lengths all $2a$.

(*3*) $1000\mathbf{e} + 1000\sqrt{3}\mathbf{n} + 2000\sqrt{3}\mathbf{u}$

$980\mathbf{e} + 980\sqrt{3}\mathbf{n} + 2000\sqrt{3}\mathbf{u}$.

(*4*) $\sin\frac{\theta}{2} = 9/\sqrt{665}$. (*5*) $\frac{1}{5}$. (*7*) Yes.

Exercise 7b

(*1*) $-\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}; \frac{2}{7}, \frac{3}{7}, -\frac{6}{7}; \frac{3}{13}, -\frac{4}{13}, \frac{12}{13}$.

(2) (*a*) If $p = \sqrt{\frac{1}{3}}$: $(p, -p, p)\,(p, p, p)\,(p, p, -p)\,(p, -p, -p)$.

(*b*) If $q = \sqrt{\frac{2}{3}}, r = \sqrt{\frac{1}{3}}$: $(q, 0, r)\,(0, q, r)\,(0, q, -r)\,(q, 0, -r)$.

(*3*) $AB(1, 0, 0)$ $BC(0, 1, 0)$ $CD(-1, 0, 0)$ $DA(0, -1, 0)$ $VA(-\frac{1}{2}, -\frac{1}{2}, \sqrt{\frac{1}{2}})$ $VB(\frac{1}{2}, -\frac{1}{2}, \sqrt{\frac{1}{2}})$ $VC(\frac{1}{2}, \frac{1}{2}, \sqrt{\frac{1}{2}})$ $VD(-\frac{1}{2}, \frac{1}{2}, \sqrt{\frac{1}{2}})$.

(*4*) $\left(0, -\frac{4}{5}, \frac{3}{5}\right)\left(\frac{3}{5}, 0, \frac{-4}{5}\right)\left(\frac{-3}{\sqrt{74}}\,\frac{-4}{\sqrt{74}}\,\frac{7}{\sqrt{74}}\right)\left(\frac{3}{\sqrt{74}}, \frac{4}{\sqrt{74}}, \frac{-7}{\sqrt{74}}\right)$

$\left(\frac{-3}{5}, 0, \frac{4}{5}\right)\left(0, \frac{4}{5}, \frac{-3}{5}\right)$

(*5*) 41, $\left(\frac{24}{41}, \frac{32}{41}, \frac{9}{41}\right)$ and $\left(\frac{4800}{41}, \frac{6400}{41}, \frac{1800}{41}\right)$

(*6*) $13°$; (*a*) $|\mathbf{H}| = 40$, $|\mathbf{V}| = 9$; (*b*) 037° to nearest degree.

Exercise 7c

(*1*) $\cos\theta = \frac{12}{13}$; $\theta \simeq 22\frac{1}{2}$. (*2*) 400 ft.

(*5*) Mark sums are equal; 'moduli' $10\sqrt{174}$, $10\sqrt{180}$.

'$\cos\phi$' $= \dfrac{17200}{|\mathbf{a}|\,|\mathbf{b}|} \simeq 0{\cdot}97$ exaggerates closeness: better to take as origin the *mean* marks of the exam, making all values of ϕ possible for pairs of candidates.

(*7*) Write $\mathbf{a} + \mathbf{b} = \mathbf{p}$ say, and proceed.

(*10*) $\overline{BH} = \mathbf{p} + \mathbf{h}$, $\overline{CH} = \mathbf{q} + \mathbf{h}$, $\overline{BC} = \mathbf{p} - \mathbf{q}$, $(\mathbf{p} + \mathbf{h}) \cdot \mathbf{q} = 0$, and so on.

(*12*) $\frac{1}{3}\sqrt{5}, \frac{1}{10}\sqrt{30}, \frac{2}{15}\sqrt{6}; \frac{1}{24}$.

Exercise 8a

(*1*) $\mathbf{x} = \begin{pmatrix} -5 \\ 1 \\ 7 \end{pmatrix} + t\begin{pmatrix} 4 \\ 2 \\ -4 \end{pmatrix}$ or $\begin{pmatrix} -5 + \frac{2}{3}s \\ 1 + \frac{1}{3}s \\ 7 - \frac{2}{3}s \end{pmatrix}$; $(-9, -1, 11)$

(*2*) Q, $(5, 6, -3)$; R, $(-2, 5/2, 4)$; $t = 3/4$, $s = 9/2$

(*3*) $(-5, 1, 7)\,(-3, 2, 5)\,(3, 5, -1)\,(-11/3, 5/3, 17/3)$; $3/7$; $AP{:}PB = 3{:}4$.

(*5*) Midpoint of BC. (*7*) $(-1, 1, 1.)$

Exercise 8b

(*2*) (4, 12, 6). (*3*) $4x - y - 3z = 7$.

(*4*) $\frac{x-2}{9} = \frac{y+1}{-2} = \frac{z+1}{-7}$. (*5*) (2, 1, 2).

(*6*) $\mathbf{x} \cdot \hat{\mathbf{n}} = 13$ where $\hat{\mathbf{n}} = \begin{pmatrix} 4/13 \\ 12/13 \\ 8/13 \end{pmatrix}$; $4x_1 + 12x_2 + 3x_3 = 169$. $\frac{169}{4}, \frac{169}{12}, \frac{169}{3}$.

(*7*) $4x_1 + 12x_2 = 169$, $x_3 = 0$; $(3/\sqrt{10}, -1/\sqrt{10}, 0)$; 3, 9, -40.

(*8*) $-7, 1, 16$; $\begin{pmatrix} -1 - 7t \\ 4 + t \\ 2 + 16t \end{pmatrix}$

(*9*) $\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}$ (all 8 combinations of signs), $\frac{a}{\sqrt{3}}$.

Exercise 8c

(*1*) (*a*) Not. (*b*) (8, 1, 2). (*2*) $3x_1 + 4x_2 + 5x_3 = 12$; $ax_1 + bx_2 + cx_3 = a + b + c$.

(*3*) (7, 8, 2). (*4*) $(2/r, -3/r, -7/r)$ or reversed, where $r = \sqrt{62}$.

(*5*) $z - 4 = \frac{1}{2}(x - 2) + (y - 1)$.

(*6*) $\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ -9 \\ 7 \end{pmatrix}$; $\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$; no.

(*7*) $2x_1 - x_2 + 2x_3 = 9$ *or* $\mathbf{x} \cdot \begin{pmatrix} 2/3 \\ -1/3 \\ 2/3 \end{pmatrix} = 3$; 2; opposite side.

Exercise 8d

(*1*) $\mathbf{x} = \begin{pmatrix} 6 + 5t \\ 5 - 3t \\ 4 - 2t \end{pmatrix}$. (*2*) $\sqrt{2}$. (*3*) $\begin{pmatrix} -1 - 3t \\ 1 + 2t \\ t \end{pmatrix}$.

Exercise 8e

(*1*) $\sqrt{(R^2 - p^2)}$, $R \sin\alpha$.

(*2*) $(a\cos t)\mathbf{i} + (a\sin t)\mathbf{j} + h\mathbf{k}$, etc., $\frac{d\mathbf{x}}{dt} = (-a\sin t)\mathbf{i} + (a\cos t)\mathbf{j}$, etc.

(*3*) $l_1 + l_2 + l_3 > 1$; $2\sqrt{2h}$, on ray with cosines 2/3, 2/3, $-1/3$.

(*4*) $x_1^2 - x_2^2 - x_3^2 = 0$; $x_2 = 1$, $x_1^2 - x_3^2 = 1$ (rect. hyperbola).

(*5*) $3x_1^2 + 4x_2^2 = 4x_1x_3 + 12x_1$.

Exercise 8f

(*1*) (*i*) $\cos^{-1}(1/3)$, (*ii*) $\cos^{-1}\dfrac{1}{\sqrt{26}}$, (*iii*) 40/3. (*2*) $x+2y = 3$. (*3*) $x-y+z = -2$.

(*4*) $\begin{pmatrix} -1 \\ 5+3s \\ 4+2s \end{pmatrix}$. (*5*) $\tan^{-1}(1/20)$, 1 mile.

(*8*) Plane is ⊥ to axis of cone and meets it in circle, i.e. where it also meets a sphere.

(*9*) (*i*) Trajectory, parabola in x_1Ox_3 plane.
(*ii*) Spiral $r = e^t$, in plane x_1Ox_2.
(*iii*) Curve on unit sphere spiralling towards pole ($t = \pi/2$).

(*10*) Internally, in $(2, -1, 0)$; $x-4y+3z-6 = 0$.

(*11*) $x = -y = z/2$; $10x-8y-9z = 0$.

(*13*) $6x-7y-2z+12 = 0$.

(*14*) (6, 3, 2). (*15*) $\cos^{-1}\left(\dfrac{4}{21}\right)$; $1-t$, $-1+10t$, $18t$.

(*16*) (*a*) $x:y:z = 11:6:7$. (*b*) 2·2, 1·2, 1·4.
(*c*) $6+11t$, $6t$, $2+7t$. (*d*) Result (*c*), with $t = 2$, i.e. 28, 12, 16.
(*e*) No solution. (*f*) $-1+11t$, $-1+6t$, $-1+7t$.

(*17*) (*i*) $k_1 = -\frac{1}{6}$.
(*iii*) $k_2 = \frac{1}{2}$, $k_3 = -\frac{1}{2}$, $k_4 = \frac{1}{6}$

$$\begin{aligned} p &= af_1+bf_2+cf_3+df_4 \\ &= -\frac{a}{6}(x-2)(x-3)(x-4)+\frac{b}{2}(x-1)(x-3)(x-4) \\ &\quad -\frac{c}{2}(x-1)(x-2)(x-4)+\frac{d}{6}(x-1)(x-2)(x-3). \end{aligned}$$

(*18*) (*a*) R does not satisfy the condition that pr_1+qr_2 is to be a member of R for all real p, q: it is only true if $p^2+q^2 = 1$. (*c*) It fails the second test, in general: it works only for multiples of a right angle, modulo 4.

(*19*) (*iv*) Associativity does not hold.

Exercise 9

(*1*) (*a*) $\begin{pmatrix} 2 & 14 \\ 2 & 18 \end{pmatrix}$, (*b*) $\begin{pmatrix} -1 \\ 7 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 8 & 12 \\ 13 & 19 \end{pmatrix}$, (*e*) (0),
(*f*) $\begin{pmatrix} 20 & 13 \\ 13 & 8 \\ 14 & 9 \end{pmatrix}$, (*g*) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, (*h*) $\begin{pmatrix} 3 & 1 & 2 \\ 2 & 3 & -5 \\ 6 & 2 & 1 \end{pmatrix}$, (*i*) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

(*2*) $\begin{pmatrix} 4 & 7 & 10 \\ -3 & -4 & -5 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 7 \\ -4 \end{pmatrix}$, $\begin{pmatrix} 7 \\ -4 \end{pmatrix}$.

(*3*) $\begin{pmatrix} 5 & 9 \\ 6 & 10 \end{pmatrix}$, $\begin{pmatrix} 1 & 4 \\ 1 & 5 \end{pmatrix}$, $\begin{pmatrix} 3 & 7 \\ 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 6 & 13 \\ 7 & 15 \end{pmatrix}$. Yes.

(*4*) (*i*) —, (*ii*) $\begin{pmatrix} 14 \\ -3 \end{pmatrix}$, (*iii*) $\begin{pmatrix} 8 \\ 3 \\ 11 \end{pmatrix}$ (*iv*) —, (*v*) —, (*vi*) $\begin{pmatrix} 7 & 6 \\ 3 & 4 \\ 4 & -18 \end{pmatrix}$.

(*5*) 2×3, 3×2; 1×3, 3×1; 3×2, 2×2; (*i*) $n = p$; (*ii*) $m\times q$.

(*6*) $\begin{pmatrix} 10 & 14 \\ 6 & 9 \end{pmatrix}$, $\begin{pmatrix} 5 & 16 \\ 4 & 14 \end{pmatrix}$. No.

(*7*) $\begin{pmatrix} 11 & 11 \\ 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 11 & 11 \\ 1 & 1 \end{pmatrix}$.

(*8*) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

(*9*) (*i*) $\begin{pmatrix} 5 & 0 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. (*ii*) $\begin{pmatrix} 3 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & -2 & 26 \end{pmatrix}\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$.

(*10*) $(px+qy+rz)$; yes; inner product.

(*11*) (y^2-4ax); a parabola.

Exercise 10a

(*1*) Mapping; many-one; not reversible.
(*2*) Not a mapping unless $x \geqslant 0$; then one-one and reversible.
(*3*) Mapping; many-one; not reversible.
(*4*) Mapping; many-one; not reversible.
(*5*) Mapping; many-one; not reversible.
(*6*) Not a mapping; not everyone has a bank.
(*7*) Mapping; many-one; not reversible.
(*8*) Mapping; many-one; not reversible.
(*9*) Not a mapping; if we restrict x so that $-1 \leqslant x \leqslant 1$ and consider principal values only, then it is one-one and reversible.
(*10*) Not a mapping; if $x > 0$, a mapping; one-one; reversible.

Exercise 10b

(*2*) $2-2x$; $1-x^2$; $(1-x)^2$; 0, 1; $2-2x^2$; $2-2x^2$; V.

(*3*) i; i; fg; $x \to -1/x$; fg = gf;

i	f	g	fg
f	i	fg	g
g	fg	i	f
fg	g	f	i .

(*4*) $l^{-1}: x \to e^x$; $x^{\frac{1}{2}} = e^{\frac{1}{2}\log x}$.

(*5*)

	i	*a*	*b*	*c*	*d*	*f*
i	*i*	*a*	*b*	*c*	*d*	*f*
a	*a*	*i*	*d*	*f*	*b*	*c*
b	*b*	*c*	*i*	*a*	*f*	*d*
c	*c*	*b*	*f*	*d*	*i*	*a*
d	*d*	*f*	*a*	*i*	*c*	*b*
f	*f*	*d*	*c*	*b*	*a*	*i*

(*6*) y^2 is the identity mapping.

(*7*) $n = 4$, $m = 3$.

Exercise 11

(*1*) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \dfrac{ax}{a-y} \end{pmatrix}$; $\mathbf{i} \to (1)$; $\mathbf{j} \to (0)$; not linear.

(*2*) (*a*), (*b*), (*d*), (*g*), (*h*) and (*i*) are linear.

(*3*) Yes. (*4*) No. (*5*) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} x-z\cot\alpha \\ y \end{pmatrix}$; yes.

(*6*) $x' = ax+by$
$y' = cx+dy$

Exercise 12a

(*2*) $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$

(*6*) These transformations are (not in order) a reflection, a rotation, 2 shears and a stretch.

Exercise 12b

(*3*) $\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}; \quad \begin{pmatrix} \cos^2\alpha+\sin^2\alpha & 0 \\ 0 & \cos^2\alpha+\sin^2\alpha \end{pmatrix}.$

(*4*) $\beta-\alpha$. (*5*) Yes.

(*6*) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}.$

Exercise 13a

(*1*) $x_1 = X_1/2 - 5X_2/6;\ x_2 = X_2/3.$

(*3*) $\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}.$ (*5*) $\begin{pmatrix} 1 & -1 & 0 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix}.$

(*6*) (*a*) $\begin{pmatrix} 1 & -1 & -2 \\ 2 & -1 & -3 \\ -6 & 3 & 10 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 1 & 0 \\ -8 & 5 & 4 \\ -2 & 1 & 1 \end{pmatrix}.$

Exercise 13b

(*1*) $x = 1, y = 4, z = 3.$

(*2*) $x = 3, y = 2, z = 1.$

(*3*) (*a*) $x = 5, y = 1, z = 0.$ (*b*) $x = -5, y = -7, z = -16.$

(*5*) (*a*) $x = 2, y = 2, z = -1.$ (*b*) $x = 1, y = \frac{1}{2}, z = 1.$

(*c*) $x = 1, y = 1, z = 1.$ (*d*) $x = 1, y = \frac{1}{3}, z = 2.$

Exercise 14a

(*1*) (*a*) -1, (*b*) -2, (*c*) 0, (*d*) 1.

(2) (*i*) -15, 15; (*ii*) -67, 67; (*iii*) 3, -3; (*iv*) 21, 42; (*v*) -7, -21.

(*4*) (*i*) -32; (*ii*) -7; (*iii*) 1; (*iv*) 48; (*v*) 20; (*vi*) -41.

(*6*) (*i*) 9; (*ii*) 9; (*iii*) 9; (*iv*) 9; (*v*) 9; (*vi*) 9.

Exercise 14b

(*1*) $\begin{pmatrix} 16 & 22 & -24 \\ -6 & 7 & 9 \\ 20 & -3 & 31 \end{pmatrix}$

(*5*) $\begin{pmatrix} 30 & 9 & -57 \\ -10 & -3 & 19 \\ 10 & 3 & -19 \end{pmatrix}$, 0.

(*6*) $\begin{pmatrix} 34 & -32 & -28 \\ -17 & 16 & 14 \\ 17 & -16 & -14 \end{pmatrix}$, 0.

(*7*) All zero.

Exercise 14c

(*1*) -8; -8; -4; -4.

Exercise 14d

(*1*) 1. (*3*) 64; 50. (*4*) -159; 39. (*5*) 1.

(*6*) (*i*) -36; (*ii*) -24; (*iii*) 54; (*iv*) 14; (*v*) 30; (*vi*) 2.

Exercise 14e

(*1*) (*b*) 19 units.

(2) 29 units.

Exercise 15a

(2) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix}.$

Exercise 15c

(*1*) (*i*) $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$, (*iii*) $\begin{pmatrix} 7 & -3 \\ -9 & 4 \end{pmatrix}.$

(2) (*i*) $\begin{pmatrix} -2 & 9 \\ -1 & 4 \end{pmatrix}$, (*ii*) $\frac{1}{2}\begin{pmatrix} 3 & -4 \\ -1 & 2 \end{pmatrix}$, (*iii*) $-\frac{1}{7}\begin{pmatrix} 3 & -2 \\ -11 & 5 \end{pmatrix}$, (*iv*) no inverse,

(*v*) $-\frac{1}{4}\begin{pmatrix} 1 & -1 \\ -7 & 3 \end{pmatrix}$, (*vi*) no inverse.

(*3*) $\dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ provided $ad-bc \neq 0$.

(*4*) (*i*) $\begin{pmatrix} 18 & -7 & 3 \\ 12 & -5 & 2 \\ 7 & -3 & 1 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 1 & 0 & -2 \\ -3 & 1 & 8 \\ -7 & 2 & 19 \end{pmatrix}$, (*iii*) no inverse, (*iv*) $\frac{1}{2}\begin{pmatrix} 4 & 2 & -8 \\ 3 & -1 & -5 \\ 1 & -1 & -1 \end{pmatrix}.$

Exercise 15d

(*1*) $\begin{pmatrix} 2{\cdot}7790 & -0{\cdot}4633 \\ -1{\cdot}1164 & 1{\cdot}8529 \end{pmatrix}.$

(*2*) $\begin{pmatrix} 3{\cdot}0433 & 0{\cdot}6957 \\ -0{\cdot}5435 & 1{\cdot}3043 \end{pmatrix}.$

Exercise 16a

(*1*) $\begin{matrix} -2 & 3 & 5 \\ 30 & -13 & -11 \\ -22 & 17 & 7 \end{matrix}$

(2) $\begin{vmatrix} 12 & -4 & 0 \\ 21 & -9 & -3 \\ -54 & 26 & 6 \end{vmatrix}, \begin{vmatrix} 24 & 36 & 60 \\ 24 & 72 & -96 \\ 12 & 36 & -24 \end{vmatrix}.$

Exercise 16b

(*1*) $\frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ -2 & -5 & 11 \\ 3 & 9 & -18 \end{pmatrix}$

(2) (*a*) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (*b*) $\frac{1}{11}\begin{pmatrix} 29 & 1 & -51 \\ -11 & 0 & 22 \\ -14 & -2 & 25 \end{pmatrix}$ (*c*) No inverse

(*d*) $1/10\begin{pmatrix} 10 & 0 & 0 \\ 0 & -1 & 4 \\ 0 & 3 & -2 \end{pmatrix}$ (*e*) No inverse (*f*) $1/180\begin{pmatrix} -51 & 48 & -45 \\ 38 & -4 & 30 \\ -24 & 12 & 0 \end{pmatrix}$.

(*4*) $\frac{1}{(1+\sigma)(1-2\sigma)}\begin{pmatrix} 1-\sigma & \sigma & \sigma \\ \sigma & 1-\sigma & \sigma \\ \sigma & \sigma & 1-\sigma \end{pmatrix}$ provided $\sigma \neq \frac{1}{2}$ or -1.

(*5*) $\begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & \sin\theta \\ \sin\theta\cos\phi & \sin\theta\sin\phi & -\cos\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix}$

Exercise 16c

(*1*) (*a*) $\frac{1}{2}\begin{pmatrix} 2 & -2 & -4 \\ 6 & -5 & -13 \\ -22 & 19 & 49 \end{pmatrix}$ (*b*) $\frac{1}{6}\begin{pmatrix} 29 & 1 & 9 \\ 31 & -1 & 9 \\ 13 & -1 & 3 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$.

(2) (*a*) $\frac{1}{2}\begin{pmatrix} -59 & -43 & 13 \\ 28 & 20 & -6 \\ -8 & -6 & 2 \end{pmatrix}$ (*b*) $\frac{1}{4}\begin{pmatrix} -30 & 22 & 14 \\ 22 & -16 & -10 \\ -2 & 4 & 2 \end{pmatrix}$ (*c*) No inverse.

(*3*) (*a*) $\frac{1}{3}\begin{pmatrix} 6 & 0 & -1 \\ 3 & -3 & -2 \\ -3 & 6 & 3 \end{pmatrix}$ (*b*) $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ (*c*) $\begin{pmatrix} 4 & -12 & 7 \\ 0 & -3 & 2 \\ 1 & -2 & 1 \end{pmatrix}$.

Exercise 17a

(*1*) $\lambda = 5$; $x; y:z = -1:1:1$.
(*2*) $\lambda = 6$ or -1; $x:y:z = 5:-2:-3$ or $1:1:-2$.
(*3*) If $\lambda = -5$, $x:y:z = 0:1:2$; otherwise all zero.
(*4*) $\lambda = -3$; 3, 7, 5.
(*5*) $\lambda = 1, 2$ or 3; $x:y:z = -1:0:1$ or $-2:1:0$ or $0:1:-1$.
(*6*) $\lambda = 2, 1$ or -1; $x:y:z = 1:1:1$ or $1:2:2$ or $1:2:3$.

Exercise 17b

(*3*) $p = 0$.

Exercise 17c

(*1*) $x = (2+ab)/(a+2)$, $y = (b-1)/(a+2)$ unless $a = -2$; if $a = -2$ and $b \neq 1$ equations are inconsistent; if $a = -2$ and $b = 1$, $x = 1-2t$, $y = t$.
(*2*) 3; $a \neq 2$, inconsistent; $a = 2$, $x = t$, $y = 2+3t$.
(*3*) $x = (4a+3b)/(a^2+3b)$, $y = (4b-ab)/(a^2+3b)$ if $a^2+3b \neq 0$; if $a^2+3b = 0$ and $a = 0$, $x = t$, $y = 4/3$; if $a^2+3b = 0$ and $a = 4$, $x = u$, $y = 4(1-u)/3$; if $a^2+3b = 0$ and $a \neq 0$ or 4, inconsistent.
(*4*) If $\lambda = 0$, $x = t$, $y = 0$; if $\lambda = 1$, $x = u$, $y = 1-u$.
(*5*) $ab-1 = 0$; if $a = 1$ and $b = 1$, $x = t$, $y = 1-t$; otherwise inconsistent.

(*6*) (*i*) If $m \neq 1$, $x = \dfrac{2-b}{1-m}$, $y = \dfrac{2-mb}{1-m}$.
(*ii*) If $m = 1$, no solution if $b \neq 2$.
(*iii*) If $m = 1$, $b = 2$, then $x = t$, $y = 2+t$.
Case (*i*) Two non-parallel lines.
Case (*ii*) Two parallel lines not intersecting.
Case (*iii*) Two coincident parallel lines.

Exercise 17d

(*1*) $k = 0$, inconsistent; $k = -4$, $x = (12+35t)/7$, $y = (77t-10)/7$, $z = 7t$.
(2) $\lambda = 7$ or -10; $(2, -1)$, $(-1, 1)$.
(*3*) $x = \frac{1}{3}$, $y = -\frac{1}{2}$, $z = \frac{5}{6}$; $x = t$, $y = (-5-3t)/12$, $z = (33t-1)/12$.
(*4*) 8. (*5*) $\lambda = 1$, $x = t$, $y = (1+5t)/8$, $z = (7t-5)/16$.
(*6*) $x = 11$, $y = 10$, $z = 12$; $x = 1+t$, $y = t$, $z = 2+t$; inconsistent.
(*7*) $b \neq 3\frac{1}{2}$; $b = 3\frac{1}{2}$. (*8*) Consistent when $b = 1$; $x = t$, $y = 1+t$, $z = -1-3t$.
(*9*) $(a+b+c)(a^2+b^2+c^2-bc-ca-ab)$; $x = (a^2-bc)/\Delta$; $y = (b^2-ca)/\Delta$; $z = (c^2-ab)/\Delta$; inconsistent.
(*10*) $x = t$, $y = -t$, $z = 0$; $x = t$, $y = -t$, $z = 0$; $x = -t$, $y = -t$, $z = 2t$.

Exercise 18a

(*1*) $\begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$, $\begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}$, $\begin{pmatrix} 14 & 9 \\ 31 & 20 \end{pmatrix}$.

(*5*) $ad-bc \neq 0$; $\dfrac{x(1+p)+(1-p)}{x(p-1)-(1+p)}$; $\dfrac{x(p+1)-(p-1)}{x(p-1)-(p+1)}$.

Exercise 18b

(*2*) Both $\mathbf{AA'}$ and $\mathbf{A'A}$ are square, with symmetry about the diagonal.
(*6*) $\mathbf{CBA'}$.

Exercise 18c

(*3*) Symmetric for m even and skew-symmetric for m odd. (*4*) Yes.

Exercise 18e

(*7*) (*a*) $p = 1, r = 0$; (*b*) $ps-qr = \pm 1$; (*c*) $p = s = k, q = r = 0$; (*d*) $p = s, q = -r$, $p^2+q^2 = 1$; (*e*) $p = 1, r = 0, s = \pm 1$. $\mathbf{K}$ is type *e*.
$\begin{pmatrix} 4 \\ 0 \end{pmatrix} \to \begin{pmatrix} 136/25 \\ 27/25 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 3 \end{pmatrix} \to \begin{pmatrix} -36/25 \\ 48/25 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 3 \end{pmatrix} \to \begin{pmatrix} 4 \\ 3 \end{pmatrix}$.
(*9*) $(bc+ca+ab)^2$.
(*10*) $(b-c)^2(c-a)^2(a-b)^2(b+c)(c+a)(a+b)$.

Exercise 18f

(*1*) $\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$.
(2) Yes.
(*3*) 2, -3; $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$; $1/\sqrt{5}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $1/\sqrt{5}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.
(*5*) $\mathbf{B}$ is skew-symmetric.

Exercise 18g

(*3*) Yes. (*7*) Skew-symmetric.

(*8*) $\begin{pmatrix} 1 & 4 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$; yes.

Exercise 18h

(*2*) Yes. (*3*) Yes. (*7*) Singular matrices. (*8*) Yes.

Exercise 18i

(*2*) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, $\begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix}$.

(*3*) No. (*4*) 4. (*5*) (*e*). (*9*) $\mathbf{I}_r = \mathbf{PAQ}$, where **P** and **Q** are non-singular.

(*11*) Determinant corresponds to discriminant.

(*13*) (*i*) $\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, (*ii*) inconsistent, (*iii*) $\begin{pmatrix} t & 2t \\ 3-2t & 6-4t \end{pmatrix}$.

Harder Examples

(*1*) Works. (*2*) $\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$, singular.

(*3*) All zero; bottom two lines have zeros.

(*4*) $5\mathbf{A} - \mathbf{I}$, $5\mathbf{I} - \mathbf{A}$.

(*5*) Yes; yes; transformation not (1 - 1).

(*6*) Direction numbers 3, 4; 4, −3.

(*7*) (*g*) Chance of being home after 8 intervals is 175/256.

(*8*) $-r$.

(*9*) Direction numbers −10, −4, 1. $dz = 10\,\partial x + 4\partial y$; $z - 10 = 10(x-2) + 4(y-1)$.

(*11*) $x = 4e^{-2t}$, $y = -4e^{-2t}$.

(*12*) $-2\mathbf{i} + 11\mathbf{j} + 16\mathbf{k}$.

(*13*) e^t, e^{2t}.

(*17*) (*i*) $x = y = z = 1/5$; (*ii*) Inconsistent; (*iii*) $x = t - \frac{2}{3}$, $y = t$, $z = t - \frac{1}{3}$;
(*iv*) $x = t$, $y = u$, $z = 1 - t - u$.

(*18*) 1:2:2; 6, 18.

(*19*) $\frac{1}{2}$, 1, $4\frac{1}{2}$; 4:4: −3; 0:2: −1; 4:4:5.

(*20*) $x = (pl - mc + nb)/\Delta$, $y = (pm - na + lc)/\Delta$, $z = (pn - lb + ma)/\Delta$, where $\Delta = l^2 + m^2 + n^2$.

(*22*) No.

BIBLIOGRAPHY

The authors have found the following books useful.

An O level course:

1. *Matrices I and II* by G. Matthews, published by Edward Arnold.

Books suitable for A level courses:

2. *Elementary Vector Algebra* by A. M. Macbeath, published by O.U.P.
3. *A Geometric Introduction to Linear Algebra* by D. Pedoe, published by Wiley.
4. *Numerical Mathematics* by A. J. Moakes, published by Macmillan.

Scholarship or first-year university texts:

5. *Algebraic Structure and Matrices* by E. A. Maxwell, published by C.U.P.
6. *Linear Equations* by P. M. Cohn, published by Routledge.

Specialized university texts:

7. *Finite Dimensional Vector Spaces* by P. R. Halmos, published by Van Nostrand.
8. *Linear Geometry* by Gruenberg and Weir, published by Van Nostrand.
9. *Determinants and Matrices* by A. C. Aitken, published by Oliver and Boyd.
10. *Introduction to Linear Transformations and Matrices* by D. Finkbeiner, published by W. H. Freeman.

General university algebra texts:

11. *A Survey of Modern Algebra* by Birkhoff and Maclane, published by Collier Macmillan.
12. *Algebra* by J. W. Archbold, published by Pitman.

INDEX

O&B H. Neill & A.J. Moakes **Vectors, Matrices & Linear Equations**